

34 spArGtYrAlAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
 101 ATGATATGCAAAACTTGAGCAATTCCTTTACTAAGCACTTACCGGT 150
 51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
 151 GTGATATTATGCGACCGAATCTTAGAAAAATCATGCTGCTAGCAGA 200
 67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuCys 84
 201 GGCTTAAAGAAATTATGGTTGGTTCATGGAAGAAATTCGCTTAGCA 250
 84 eRcGluAsnGlyGluGluPhePheIleProValLeuAlaGlyLeuPhe 100
 251 GTGAAACCTGTAAGAAATCTTATCTCTATAGCCGCTTATATTATA 300
 101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgLysLeuVal 117
 301 GGTCGCGTGTGGCTCCAACTAATGATTTACCTAGCTAGGAAATGGT 350
 117 LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
 351 TCACAGATTAGGCACCTCTAAGCAACATGTGATTTAGTCTAAAAAG 400
 134 LysLeuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150
 401 GATTAGATTAAGTTAATCTTACCAAAAAACGTAACGCTATTAAAAAC 450
 151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
 451 ATGTGATATTGACAGCAAGTGAATTAAGGTTATCAATTCATGCA 500
 167 pAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
 501 CAACCTTATTAATAAAAAACATCCCAAGGTTCAAGATCAAGTATTTA 550
 184 ySThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
 551 AAACGTGAAAGTTAACCGCAAGAACAAAGTTGCTTATTAATGAACCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGlyAsnIle 217
 601 TCGGCTCAACCGGTTGCCAAAAGGTGCAACTTACTCATGAATAATT 650
 217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
 651 GGTCACTGATTTTCTCAAGCTAGAGATCAATTTATGGAACCAACATTT 700
 234 ePProGlyThrAlaIleLeuThrValAlaProPheHisHisGlyPheGly 250
 701 CACCGAGGACGGCTATTTTAAGTCAATGATCCATCCATCATGCTTTTGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIle 267
 751 ATGTTTACTACTTAAAGCTATCTAACTGTGTTTCTGATTTGCATGCTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrIleS 284
 801 AACGAAATTTGACAGAGAGACTTTTAAAAAACCTGCAAGATTACAAAT 850
 284 ySSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 851 GTTCAAGGTTATCTTGTACCGACTTTGTTGCAATTCCTTAATGAGAGT 900
 301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGly 317
 901 GAATTACTCGATTAATATGATTTATCAAAATTTAGTGAATTCATCTCG 950
 317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
 951 CGGAGCCTCTTATCTAAGAAATTTGGAAGCTGTGCTAGAGCTTTTA 1000

334 snLeuProGlyValAlaArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
 1001 ATTTACCGGGTGTCTGTCACAGCTATGGTTTAAACAGAAACACTCTGCA 1050
 351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVal 367
 1051 ATTATTATCACACCGAAGGAGATGATTAACACAGCTCTCTGGCAAGT 1100
 367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThr 384
 1101 TGTGCTATTATTAAACCAAAAGTTATCGATCTTGAATCTTAAAAAACTT 1150
 384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
 1151 TGGGCCCGAAGACAGCGTGGAAGATTGTGTAAAGGCTCTATGCTTAGT 1200
 401 LysGlyTyrValAlaAspAsnProGluAlaThrArgGluIleIleAspGlu 417
 1201 AAAGTTATGATAGATTAATCCAGAAACAAAGAAATCATAGATGAGA 1250
 417 uGlyThrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
 1251 AGTTGGTTCACACACAGAGATATTGGTATTTACGATGAGAAAGAAACAT 1300
 434 hePheIleValAlaAspArgLeuLysSerLeuIleLysTyrGlyTyrGln 450
 1301 TCTTTATCGTGAGATCGTTGAAGCTTATCAAAATCAAGAAATTCACA 1350
 451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
 1351 GTACACCTCTGTAATTAAGAAATCTGTTTTCACACATCCAAATATTTT 1400
 467 eAspAlaGlyValAlaAlaGlyValProAspProIleAlaGlyLeuProG 484
 1401 TGATGCCGGGTTGTCGGGCTTCACAGTCTTATAGCTGGAGCTTCGG 1450
 484 lYalAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
 1451 GAGCTGTTGTTGTAATTAAAGAAAGAAATCTATGACTGAAAGAAAGTA 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuAlaGlyL 517
 1501 ATGATTTACTTCTGCTACTCAAGTTCAAAATGCAAAAGCTTGGCGTGTG 1550
 517 yValArgPheValAlaAspGluValProLysGlyLeuThrGlyLysIleAsp 534
 1551 TGTCCGTTTGTGACGAGAGTACTTAAGGTCTCACTGGTTAAATTTGAG 1600
 534 lYysAlaIleLeuArgGluIleLeuLysLysProValAlaLysMet 548
 1601 GTAAGCAATTAGAGAAATTAAGTAAACCAAGTGTCTAAGATG 1644
 seq_name: /SIDS2/gcdata/geneseq/geneseq/NM1999.DAT:AAx84396
 seq_documentation_block:
 ID AAx84396 standard; DNA; 1644 BP.
 XX
 AC AAx84396;
 XX
 DT 09-SEP-1999 (first entry)
 XX
 DE L. lateralis luciferase coding sequence.
 XX
 KW Luciferase; surfactant-tolerant; firefly; intracellular ATP assay; ss.
 OS Luciola lateralis.
 XX
 PN W09933997-A1.
 XX
 PD 08-JUL-1999.
 XX
 PF 24-DEC-1998; 98WO-JP05864.
 XX

PR 26-DEC-1997; 97JP-0361022.
 PA (KIRK) KIRKMAN CORP.
 XX Hattori N, Murakami S;
 XX WPI: 1999-419109/35.
 DR P-PSDB: AAY22183.
 PT Modified luciferase tolerant to surfactants and useful for assay of
 XX Intracellular ATP
 XX Example 2; Page 43-48; 56pp; English.
 XX This sequence encodes a luciola lateralis luciferase of the invention.
 CC The protein is a surfactant-tolerant luciferase which is derived from
 CC firefly luciferase by replacement of the glutamic acid residue at
 CC position 490 by another amino acid. The luciferase can be used in an
 CC assay of intracellular ATP in biological samples containing cells, for
 CC example for establishing the presence of foodstuffs, drugs,
 CC agrochemicals, beverages and environmental samples. The activity of the
 CC surfactant-tolerant luciferase is not lowered by the presence of high
 CC concentrations of surfactant and therefore an assay of high accuracy can
 CC be carried out in a simple manner.
 XX SQ Sequence 1644 BP; 529 A; 262 C; 348 G; 505 T; 0 other:

alignment_scores:
 Quality: 2821.00 Length: 548
 Ratio: 5.148 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.818

alignment_block:
 US-09-581-241-4 x AAY84396 ..

Align seg 1/1 to: AAY84396 from: 1 to: 1644

```

1 MetGluSnmMetGluSnmSpgUaSnIleValTYrGlyProGluProh 17
1 ATGGAAAACATGAGACGATGAAATATTGTGTATGCTGCAACCAT 50
17 eTYrProIleGluGlySerAlaGlyAlaGluLeuArgLysTYrMetA 34
51 TTACCCATTTCAGAGAGGATCTGCTGGAGCACAATTGCGCAAGTATATG 100
34 sPArTYrAlaLysLeuGlyAlaIleAlaIlePheThrAsnAlaLeuThrgly 50
101 ATCGATATGCAAAACCTGGAGCAATTCCTTTACTAAAGCAGCTTACCGGT 150
51 ValAspTYrThrTYrAlaGluTYrLeuGluLysSerCysGlyLeuGlyGI 67
151 GTGCAATATACGTACGCCCAATCTAGAAAAATCATCATGCTGTAGGAGA 200
67 uAlaLeuLysAsnTYrGlyLeuValValAspGlyArgIleAlaLeuCysS 84
201 GCGCTTAAAGAAATTAAGTTGGTTGTTGTTGATGGAAGAAATTCGTTATGCA 250
84 erGluSnmCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAAACGTGCAAGAAATCTTATTCCTGATATAGCCGGTTTATTTATA 300
101 GlyAlaGlyAlaIleProThrAsnGluIleTYrThrLeuArgGluLeuVa 117
301 GGTGTCGTTGGCTCCCAACTAATGATTTACACTACCTACGTAATGCTGT 350
117 lHisSerLeuGlyIleSerLysProThrIleValAlaPheSerSerLysL 134
351 TCACAGTTTAGGCATCTCTAAGCCAAACATGATTATTAGTTCTTAAGAA 400
134 lYleuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150

```

```

401 GATTAGATAAAGTTATTAATCTGTACAAAAACGGTAACCTATTAACCC 450
151 lIleValIleLeuAspSerLysValAspTYrArgGlyTYrGlnSerMetAs 167
451 ATTGTTATATTGGACACCAAGATGATTAAGAGTTATCAATTCATGCA 500
167 pAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
501 CACTTTATTAATAAAAAACACTCCACAAGGTTTCAAAAGCATCAAGTTTA 550
184 ySThrValGluValAsnArgLysGluValAlaLeuIleMetAsnSer 200
551 AACTGTAGAAATTAACCCCAAGACAAAGTGTCTTATTAATGAACCTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGluLeuThrHisGluSnm 217
601 TCGGGTTCAACCGGTTTCCCAAAAGGTGTGCACTTACTCATGAAAAAT 650
217 uValThrArgPheSerHisAlaArgAspProIleTYrGlyAsnGlnValS 234
651 CGTCACTAGATTTTCTCAGCTAGAGATCCAAATTAAGAAACCAAGTTT 700
234 erProGlyThrAlaIleLeuThrValAlaProPheHisGlyPheGly 250
701 CACCAGGACAGCGCTATTTAAGTATGACATTCATCATCATGTTTGGT 750
251 MetPheThrLeuGlyTYrLeuThrCysGlyPheArgIleValMetIle 267
751 ATGTTTACTTACTTAGGCTATTAACCTGTGTTTCCGATTTGATATGT 800
267 uThrLysPheAspGluLysThrPheLeuLysThrLeuGlnAspTYrLysC 284
801 AACGAAATTTGACGAAGACATTTTAAAAACACGCAAGTTCACAAAT 850
284 ySSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAAGCGTTATTTCTGTACCGACTTGTGTTCATTCCTTAATGAAAGT 900
301 GluLeuLeuAspLysTYrAspLysSerAsnLeuValGluIleAlaSerGI 317
901 GAATTAATCTCGATTAATAATGATTTATCAAAATTTACTTCAAAATGCACTCG 950
317 yGlyAlaProLeuSerLysGluIleGlyAlaValAlaArgArgPheA 334
951 CCGAGCACCCTTTATCTAAAGAAATGGAGAGCTGTGTAGAGTTTA 1000
334 sLeuProGlyValArgGlnGlyTYrGlyLeuThrGluThrSerAla 350
1001 ATTTACCGGGGTTCGTCAAGGCTATGTTTAAAGAAACACCTTGCA 1050
351 lIleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
1051 ATTATTAATCACACCGAAGCGATATAAACAGGTGCTTGGCCAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
1101 TGTGCCATTAATTAAGCAAAAGTTATTCATCTTCTAATAAAAAACTT 1150
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
1151 TGGGCCCGAAGACAGCGTGGAGAGTTTGTGTAAAGGTCCTATGCTTATG 1200
401 LysGlyTYrValAspAsnProGluAlaThrArgGluIleIleAspGluGI 417
1201 AAAGCTTATGTAGATTAATCCAGAGCAACAAGAAATCATATGATGAGA 1250
417 uGlyThrPheLysThrGlyAspIleGlyTYrTYrAspGluGluLysHisP 434
1251 AGTTTGTTTCCACACAGAGATATTGGGTATTAAGTAAAGAAACATTT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTYrLysGlyTYrGln 450
1301 TCTTTATCTGTGATCGTTTGAAGTCTTTAATCAATTAACAAAGATTAACA 1350

```



```

|||||
751 ATGTTACTACTTGTAGGCTATCTACTGTTGCTTCTATTCATCTT 800
267 urhrlyspheaspjluglurhrpheeluysphrleuglnaspyrlysc 284
801 AACGAATTGACGAGAGACTTTTAAACACTGCACAGATTACAAT 850
284 ysserSeraIleuValProthrleuphealaleuasnrgser 300
851 GTTCAGCGGTATCTTGTACGACTTCTTGCATTTCAATTTAATGAAGT 900
301 GluleuLeuaspjlyrtyraspleuserasneleVallealeSergI 317
901 GAATTACTCGATAAATATGATTTATCAATTTAGTTGAATTCGATCTGCG 950
317 yglYalaProleuserlysgluileglylualeValalaaghrphea 334
951 CGGAGCACCTTTATCTAAAGAAATGCTGAAGCTGCTGACGCTTTTA 1000
334 snleupProglyValArggluglytyrGlyleuthrgluhrhrSeraIa 350
1001 ATTTACCGGGTTCGTCAGAGGCTATGCTTTAACAGAAACACCTCTGCA 1050
351 lleilelethrProgluglyaspasplysProglyAlaSerGlytyrVa 367
1051 ATTATTTACACACCGAGGCGATGATAAACAGGCTCTTCTGCAAAAGT 1100
367 lYalProleupheulysAlalyValIleaspleuaspThrlystThul 384
1101 TGTGCCATTATTTAAAGCAAAAGTTATCGATCTTGATACATAAAACCT 1150
384 euGIYProasnArgrgllygluValCysValLysGlyPrometleumet 400
1151 TGGGCCCGACAGACGCTGAGAGGTTGCTAAAGGGCTCTATCTTATG 1200
401 LysGlyTYrValAspAsnProgluAlaThrArggluileleaspgJugI 417
1201 AAAGGTTATGATGATATTCACAGAACCAACAGAAATCATAGATGAAGA 1250
417 uGIYrPLeuHlstrnglyaspIleGlyTYrTYrAspGlucIuysHlSP 434
1251 AGGTGTGTCACACAGAGATATGGGATATACGATGACAGAAAAACATT 1300
434 hepHeileValAsparGleuysSerleuIleLysTYrLysGlyTYrGln 450
1301 TCTTATCGTGATCGTTGAAGTCTTTATCAATACAAAGATATCA 1350
451 ValProProAlaGluleuGluserValleuLeuGlnHlSPProAsnIlePh 467
1351 GTTACACCTCTGATTAAGATCTGTTTGCACATCCAAATATTTT 1400
467 eaSPAlaGlyValAlaGlyValProaspProIlealaglYgluLeuProG 484
1401 TGATGCCGGGCTGCTGGCGTCCAGATCTTAAGCTGCTGACCTCCGG 1450
484 lYalaValleValleuLysLysGlyLysSerMetThrGluysGluVal 500
1451 GAGCTGTGTTGTTACTTAAGAAAGAAATCTATGACTGAAAAAGAAAGA 1500
501 MetaspTYrValAlaSerGlnValSerAsnAlaLysArGleuArGlygl 517
1501 ATGATTCAGTTCAGTCAAGTTCAATGCAAAACGTTTGGCTGCTGCG 1550
517 yValArpHeValaspGluValProLysGlyLeuThrGlyLysIleaspG 534
1551 TGTCCGTTTTGTGACGAGACTACTAAAGCTCCTACGTGTAATATGACG 1600
534 lYlYsAlaIleArggluileuLysLysProValAlaLysMet 548
1601 GTAAAGCAATTAGAGAAATACTGAAGAAACAGTTGCTAAGATG 1644

```

seq_name: /SIDS2/gcdata/geneseq/geneseqn/NA1993.DAT.AA034745

seq_documentation_block:

ID AA034745 standard; cDNA to mRNA; 1644 BP.

AC AA034745;

DT 24-MAY-1993 (first entry)

DE Thermostable luciferase (wild-type).

KW Firefly; Luciola cruciata; GENJIT; Luciola lateralis; HEIKE;

OS mutant; assay; ss.

PN Luciola lateralis.

PD EP524448-A.

PN 27-JAN-1993.

PF 26-JUN-1992; 92EP-0110808.

PR 27-JUN-1991; 91JP-0157117.

PR 29-NOV-1991; 91JP-0317064.

PR 22-MAY-1992; 92JP-0131057.

PA (KIKR) KIKROMAN CORP.

PI Elchil N, Naoki K;

DR WPI: 1993-028553/04.

DR P-PSDB: AAR30803.

PT New DNA sequence of a thermostable luciferase - has the amino acid

PT at position-217 of Luciola cruciata or lateralis luciferase

PT replaced by a hydrophobic amino acid

PS Claim 1-3; Page 29-30 + 21-24; 33pp; English.

XX A DNA sequence of a thermostable luciferase of a firefly, which

CC encodes the amino acid sequence of a wild-type firefly luciferase in

CC which an amino acid at the 217 position or an amino acid at the

CC position equiv. to the 217 position of the luciferase of Luciola

CC cruciata (GENJIT firefly) of Luciola lateralis (HEIKE firefly) is

CC replaced by a hydrophobic amino acids, esp. Ile, Leu or Val, is

CC claimed. The mutant luciferase is identical in properties to the

CC wild-type luciferase except that it is stable when heated to high

CC temps., e.g. 50 degrees C. The luciferase can be used in assays for

CC e.g. ATP.

SQ Sequence 1644 BP; 529 A; 262 C; 349 G; 504 T; 0 other;

alignment_scores:

Quality: 2814.00 Length: 548

Ratio: 5.144 Gaps: 0

Percent Similarity: 99.818 Percent Identity: 99.635

alignment_block:

US-09-581-241-4 x AA034745 ..

Align seg 1/1 to: AA034745 from: 1 to: 1644

1 MetGlnsnMetGluasnAspGluasnIleValTYrGlyProgluProh 17

1 ATGGAAACATGAGAGAGATGAGAAATATTTGTTGTTGCTTCAACCAT 50

17 eTYrProIleGluGluGlySerAlaGlyAlaGlnLeuArGlyTyrMet 34

51 TTACCTATTGAGAGAGATGCTGAGCACAATTCGCCAAGTATATG 100

34 spArgTYrAlaLysLeuGlyAlaIlealaphethrrasAlaLeuThrGly 50

101 ATCGATATGCAAAACCTTGAGCAATTCCTTTACTAAGCATTACCGGT 150

[illegible]

```

FT CDS 1648..1908
FT /tag=
FT /note= "encodes biotinylated peptide portion of fusion
FT protein"
PN JP08308578-A.
XX 26-NOV-1996.
XX 24-APR-1995; 95JP-0039857.
XX 14-MAR-1995; 95JP-0054625.
XX 27-JUL-1994; 94JP-0193798.
XX (KIKK ) KIKKOMAN CORP.
XX WPI; 1997-059697/06.
XX P-PSDB; AAM12395.
DR Fusion protein comprising firefly luciferase and biotinylated
DR peptide - useful in a bio-luminescent analytical method for
XX quantifying ligands
XX
XX Example 7; Page 11-12; 13pp; Japanese.
XX
XX This is the nucleotide sequence encoding a novel fusion protein which
XX comprises the firefly (Luciola lateralis) luciferase protein and an 87
XX amino acid biotinylated peptide designated the E.coli biotin carboxy
XX carrier protein (BCCP-87). The fusion gene was generated by firstly
XX amplifying the sequence encoding the E.coli (BCCP-87) from the E.coli
XX genome and inserting the resultant sequence into the 3' end of the
XX firefly luciferase gene in plasmid pHLf230. The biotin peptide sequence
XX replaces the C-terminal Met residue of the wild type luciferase sequence.
XX The plasmid was then introduced into E.coli JM101 for production of the
XX fusion protein. The novel protein can be used in bioluminescence assays
XX to quantify luciferase ligands which may modulate binding of luciferase
XX to its receptor.
XX
XX Sequence 1908 BP; 598 A; 329 C; 426 G; 555 T; 0 other:
XX
XX alignment_scores:
XX Quality: 2814.00 Length: 547
XX Ratio: 5.144 Gaps: 0
XX Percent Similarity: 100.000 Percent Identity: 99.817
XX
XX alignment_block:
XX US-09-581-241-4 x AAT63269 ..
XX
XX Align seg 1/1 to: AAT63269 from: 1 to: 1908
XX
XX 1 MetGluAsnMetGluAsnAspGluAsnIleValItyrGlyProGluProPh 17
XX 1 ATGGAAACATGAGAGACATGAAATATGTTGTTAGTGGCTGGAACCATTT 50
XX 17 eTyrProIleGluGluGlySerAlaGlyValAlaGluLeuArgLysTyrMetA 34
XX 51 TTACCTTATTGGAAGAGGATCTGCTGAGCACAAATTGCCAAGTATATGG 100
XX 34 sPArGTYrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
XX 101 ATCGATATGCAAAACTTGGAGCAATTCCTTTACTAACGACCTTACCGGT 150
XX 51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
XX 151 GTGGATTATACGCGCAATCTTAGAAAAATCATGCTGCTAGGAGA 200
XX 67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
XX 201 GGCTTTAAAGATTTATGTTGGTTGTGATGGAAGAAATTCCTATGCA 250
XX 84 eArgLysCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100

```

```

251 GTGAAACCTGGAAGAGTCTTTATTCCTGATATAGCCGGTTATTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeu 117
301 GGTGCGGTGTGGCTCCAACTAATGAGATTTTACCTACTGTAATGGT 350
117 LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
351 TCACAGTTTAGGCATCTTACCCACCAATGTTATTTAGTTCTAAAAAG 400
134 lYLeuAspLysValIleThrValGlnLysThrValAlaIleLysThr 150
401 GATTAGATAAAGTTATTAAGTACAAAAAGGTTCAACTGCTATTTAAACC 450
151 lIleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
451 ATTGTTATATTGGACAGCAAAAGTGATTTATGAGTTTATCAATCATGGA 500
167 pAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerPheL 184
501 CAACCTTATTAAAAAACAACCTCCACAAGGTTTCAAGGATCAAGTTTGA 550
184 ySThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAACCTGTAAGATTAAACCGCAAAAGCAAGTGTCTTATTAATGAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
601 TCGGGTTCAACCGGTTGGCCAAAAGGTGCAACTTACTCATGAAATTT 650
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
651 GGTACAGCGGTTTTTTCACGCTAGAGATCCAAATTATGGAACCAAGTT 700
234 eArgProGlyThrAlaIleLeuThrValValProPheHisGlyPheGly 250
701 CACCAAGCAGCGCTATTATTACTGTAGTACCATTCATCATGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMet 267
751 ATGTTTACTACTTAGGCTATTAACCTTGGTTTCTGATGTCATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGluAspTyrLys 284
801 AACCAAAATTTGACGAAAGACCTTTTAAACACGCAAGATTTACAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAAGCGTATTCTTGTACCGACCTTGTTCGAATCTTAATAGAAAT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSer 317
901 GAATTACTCGATTAATATGATTTATCAAAATTTAGTTCAAAATTCAT 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgPheA 334
951 CGAGACACCTTTATCTAAGAAATTTGGTGAAGCTGTTCGACGTTTGA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSer 350
1001 ATTTACCGGTTGTCGTCAAGGCTATGTTTAAAGAAACAACCTGTGA 1050
351 lIleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLys 367
1051 ATTATTATTCACCGCGAAGGCGATATTAACACAGTGCTTGGCAAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThr 384
1101 TGTGCATTATTAAAGCAAAAGTTATCGACTTGTATCTAAAAAACTT 1150
384 eGlyProAsnArgArgGlyGluValLysValLysGlyProMetLeuMet 400
1151 TGGGCCCGAAGACAGCTGAGAAAGTTTGTTAAAGGCTCATGCTTATG 1200

```

```

401 LysGlyTyrValAspAsnProGluAlaThrArgIleIleAspGlu 417
1201 AAGGTATGATGATATATCCAGACACAGAAATCATGATGAGGA 1250
417 UGlyTrpLeuHisThrGlyAspIleGlyTyrTyrAspGluIlyHisP 434
1251 AGGTGGTTGGACACAGAGATATGGTATATGATGAGAAACATTT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1350
451 ValProPheAlaGluLeuGluSerValLeuLeuGlnHisProAsnIleP 467
1351 GTACCACTGCTGAATATGATGATGATGATGATGATGATGATGAT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyIleLeuProG 484
1401 TGATGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1450
484 lYAlaValAlaValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
1451 GAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
1501 ATGATTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
1551 TGTCGCTTTGTGTGACGACGACGACGACGACGACGACGACGACG 1600
534 lYlYsAlaIleArgGluIleLeuLysLysProValAlaLys 547
1601 GTAAAGCAATTAGGAATATCTAGAGAAACCAAGTGTCTAG 1641

seq_name: /SID52/gc9data/geneseq/geneseq/NA1998.DAT:AAV32467
seq_documentation_block:
ID AAV32467 standard; DNA: 1920 BP.
AC AAV32467;
XX
DT 10-SEP-1998 (first entry)
XX
DE Luciferase-lysostaphin fused gene.
XX
KW Luciferase-lysostaphin fused gene; chimeric; bioluminescence; ds.
OS Chimeric - Staphylococcus simulans.
OS Chimeric - Luciola lateralis.
XX
FH Key Location/Qualifiers
FT CDS 1..1920
FT /tag= a
FT /product= "Luciferase-lysostaphin fusion protein"
FT /note= "Nucleotides 1-1647 encode luciferase
FT protein while nucleotides 1648-1920 encode
FT lysostaphin; CDS does not contain stop codon"
XX
XX JP10150991-A.
XX
XX 09-JUN-1998.
XX
XX 25-NOV-1996; 96JP-0328042.
XX
XX 25-NOV-1996; 96JP-0328042.
XX
XX (KIKK ) KIKKOMAN CORP.
XX
XX WPI; 1998-379994/33.
XX
XX P-PSDB; AAW48856.
XX
DR

```

```

XX
XX New Luciferase-lysostaphin fused protein - useful for
XX bioluminescence analysis
XX
XX PS Disclosure; Pages 6-7; 10pp: Japanese.
XX
XX CC The invention claims for a luciferase-lysostaphin fused gene which
XX encodes a fused protein in which a peptide part consisting of
XX luciferase, from Luciola lateralis, is connected to a peptide part
XX consisting of lysostaphin from Staphylococcus simulans. The method
XX described by the invention can be used to prepare a luciferase-
XX lysostaphin fused protein efficiently. The fused protein is useful
XX for bioluminescence analysis.
XX
XX SQ Sequence 1920 BP; 627 A; 310 C; 407 G; 576 T; 0 other;

alignment_scores:
Quality: 2814.00 Length: 547
Ratio: 5.144 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.817

alignment_block:
US-09-581-241-4 x AAV32467
Align seg 1/1 to: AAV32467 from: 1 to: 1920

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
1 AVGGAAACATGAGACAGATGAGAAATATGATGATGATGATGATGAT 50
17 eTyrProIleGluGluGluGluSerAlaGlyAlaGluLeuArgLysTyrMet 34
51 TTACCTTATGACAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 100
34 sPARGTyrAlaLysLeuGluValAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 ATCGATATGCAAAACTTGACCAATGCTTTACTTAACGCACTTACCGGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGly 67
151 GTCCATTATACGTAGCGGCAATGCTTACCAAAATCAATGCTGTAGAGA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
201 GGCCTTAAAGAAATATGATGTTGTTGTTGATGAGAGATTCGTTATGCA 250
84 eArgLysCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAACTGTGAAGAGTCTTATCTCTGATATGACCGGTTATTATTA 300
101 GluValGlyValAlaProThrAsnGluIleTyrThrLeuArgLysVal 117
301 GGTCGTGGTGGCTCCCACTAATGACATTTACACTTACGTAATATGAT 350
117 LHisSerLeuGlyIleSerLysProThrIleValIlePheSerSerLysG 134
351 TCACAGTTTAGCATCTCTAAGCCAAATGTAATTAGTTCAATAAAG 400
134 lYLeuAspLysValIleThrValGluLysThrValAlaIleLysThr 150
401 GATTAGTAAAGTTATTAAGTACCAAAAAACGTAATCTCTATTAAACC 450
151 lIleValIleLeuAspSerLysValAspTyrArgLysTyrGlnSerMet 167
451 ATGTATATATGACAGCAAGTGATTTAGAGTTATCAATTCATGGA 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
501 CAACTTATTAATAAAACACCTCCACAAAGTTCAAAAGATCAAGTTTA 550
184 ySThrValGluValAsnArgLysGluValAlaIleLeuIleMetAsnSer 200

```

```

551 AAACGTGAGAGTTAAACCGCAAGCAAGTTGCTTATTAATGAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLe 217
    |||||
601 TCGGGTTCAACCGGTTTGGCAAAAGGTGTGCAAACTTACTCATGAAATTT 650
217 uValThrArgPheSerHisAlaArgAspIleTyrGlyAsnGlnValS 234
    |||||
651 GGTCACGCGCTTTTTCACGCTAGAGATCAATTTATGGAACCAAGTTT 700
234 expProGlyThrAlaIleLeuThrValValProPheHisGlyPheGly 250
    |||||
701 CACCAGGCGAGCTATTATTAAGTGTACATTCATCATGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
    |||||
751 ATGTTACTACTTATAGCTATCACTGATGCTGTTTCTGATTTGTCATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
    |||||
801 AACGAAATTTGACGAAGAGACTTTTAAACACACAGATTACAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
    |||||
851 GTTCAACGCTTATCTTGTACGACATTTGTTGCCAATCTTATAGAACT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
    |||||
901 GAATTACGTCATTAATGATTTATCAAAATTTAGTTGAAATTCATCTGCG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
    |||||
951 CGGAGCACCTTTATCTTAAGAAATGGTGAAGCTGTTCAGACGTTTAA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAla 350
    |||||
1001 ATTACCGGGTTCGTCAGAGCTATGTTTAAACAGAAACACCTCTGCA 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyValSerGlyLysVa 367
    |||||
1051 ATTATTTATCACCGGAGGCGAGATTAACCGAGTCTTCTGCGAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
    |||||
1101 TGTGCCATTTATTAAAGCAAAAGTTATCGATCTTGATCAAAAAAAGCT 1150
384 euGlyProAsnArgArgGlyValCysValLysGlyPrometLeuMet 400
    |||||
1151 TGGGCCCGAAGACAGCTGGAGAAAGTTGTGTAAGGGTCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgLutIleIleAspGluG 417
    |||||
1201 AAAAGTTTATGATGATATCCAGAAAGCAACAGAAATCATAGATGAGAA 1250
417 uGlyTyrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
    |||||
1251 AGGTTGGTTGACACAGAGATATTGGTATTACGATGAGAAAGAAACATT 1300
434 hepHeIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
    |||||
1301 TCTTTATGCTGATCGTTTGAAGTCTTTAATCAATACAAAGATATCTCAA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
    |||||
1351 GTACACCTGCTGATTAATGATCTGTTTGTTCACATCAAAATATTTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLutLeuProG 484
    |||||
1401 TGATCCGCGGCTTGGCTGCGTTCGATCTATACCTGCTGAGCTTCGCG 1450
484 lValAlaValValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
    |||||
1451 GAGCTGTGTTGTTACTTGAAGAAAGAAATATCTATACAGAAAAAGAACTA 1500

```

```

501 MetAspTyrValAlaIleSerGlnValSerAsnAlaLysArgLeuArgLysI 517
    |||||
1501 ATGGATTACGTTGCTAGTCAAGTTTCAATCCAAACGTTTGGCTGGTGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
    |||||
1551 TGTCCGTTTGTGGAGAGAGTACCTAAAGGTCCTCAGTGAATTTGACG 1600
534 lLysAlaIleArgGluIleLeuLysLysProValAlaLys 547
    |||||
1601 GTAAAGCAATTAGAGAAATTAAGTGAAGAAACCAAGTTGCTTAAAG 1641
seq_name: /SIDIS2/gcgdate/geneseq/geneseqn/NA1990.DAT.AA003257
seq_documentation_block:
ID AA003257 standard; DNA: 1644 BP.

```

```

XX AA003257:
AC
XX
XX 22-JUL-1990 (first entry)
DE
XX
XX Luciola lateralis recombinant luciferase gene.
XX Luciferase; enzyme; ATP; plasmid pHLf7; firefly; ss.
XX
XX Luciola lateralis.
XX
XX Key Location/Qualifiers
XX CDS 1..1644
    /*tag= a
    /product=luciferase
    /EC_number=EC-1.13.12.7

```

```

XX EP353464-A.
XX
XX PD 07-FEB-1990.
XX
XX PF 30-JUN-1989; 89EP-0111958.
XX
XX PR 22-DEC-1988; 88JP-0162402.
XX PR 01-JUL-1988; 88JP-0162402.
XX
XX PA (KIKK ) KIKKOMAN CORP.
XX
XX PI Tatsumi H, Kajiyama N, Nakano E;
XX DR WPI: 1990-038240/06.
XX DR P-PSDB; AAR03731.
XX
XX PT New gene encoding luciferase -
XX derived from Luciola lateralis and expressed in and purified
XX from Escherichia coli genus cells.
XX
XX PS Claim 3; page 24; 43pp; English.
XX
XX CC This enzyme is useful for determining ATP levels. It is
XX contained in plasmid pHLf7.
XX
XX SQ Sequence 1644 BP; 529 A; 264 C; 348 G; 503 T; 0 other;

```

```

alignment_scores:
    Quality: 2811.00      Length: 548
    Ratio: 5.139
    Percent Similarity: 99.818      Percent Identity: 99.453

```

alignment_block:

US-09-581-241-4 x AA003257 ..

Align seq 1/1 to: AA003257 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17

```

|||||
1 ATGGAAAAATGAGACGATGAAAAATATGCTATGCTCGTCAACCACT 50
17 eYrProIleGluGluGlySerIleValGluIleuArgLysTyrMet 34
|||||
51 TTACCTATGTAAGAGGATCTGCTGAGACCAATTCGCAAGTATATGG 100
34 sPaqTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
|||||
101 ATCGATTACGAAAACTTGAGCAATTCCTTACTTAACGCACTTAACCGGT 150
51 ValAspTyrThrTyrAlaGluTyrIleuGluLysSerCysLeuGlyGly 67
|||||
131 GTCGATTACGTACGCGCAATCTTACAAAAATCATGCTGTACAGAGA 200
67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuGly 84
|||||
201 GGCCTTAAGAATATATGTTGGTTGGTTGATGGAAGAAATTCGTTATGCA 250
84 eRgIAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
|||||
251 GTGAAACTGTGAAGAAATTCCTTATTCCTGTATTAGCCGGTTATTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrIleuArgGlyLeuVal 117
|||||
301 GGTGCGGTGCTGCTCCACTAATGATTAACCTCTACGTGAATGGT 350
117 hHisLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
|||||
351 TCACAGTTAGGCATCTCTAAGCCACAAATGTATTATTAAGTTCTAAGAA 400
134 LysLeuAspLysValIleThrValGluLysThrValIleThrAlaIleLysThr 150
|||||
401 GATTAGATTAAGTATTAACGTGTACAAAAAGGTAAGTCTATTAAAAACC 450
151 lLeValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
|||||
451 ATTGTATATGAGACAGAAAGTGATTAAGAGTATATCAATCCATGCA 500
167 pAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
|||||
501 CAACCTTATTAATAAAAAAACTCCACAAAGTTTCAAAAGGATCAAGTTTGA 550
184 yStrValGluValAsnArgLysGluGluValAlaLeuIleMetAsnSer 200
|||||
551 AAATGTAGAAATTAACCGCAAGAACAAAGTGTCTTATTAATGAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
|||||
601 TCGGTTCAACCGGTTGCCAAAAAGTGTCACACTTACATCATGAAATGC 650
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
|||||
651 AGTCACATGATTTTCTCAGCTAGAGATCCAAATTTATGGAACCAAGTTT 700
234 eRProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
|||||
701 CACACAGCAGCGCTATTTTAACGTAGTACCAATTCATCATGTTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgGlyLeuMet 267
|||||
751 ATGTTTACTACTTATAGCTATTAACCTTGTGTTTCTGATTGTCTATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGluAsnTyrLysC 284
|||||
801 AACGAATTTGACGAAGAGACTTTTAAAAACACCTGCAGATTTACAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
|||||
851 GTTCAAGCGCTATTTCTGTACGACTTGTGTGCAATCTTAAATGAACT 900
301 GluLeuLeuAspLysTyrThrLeuSerAsnLeuValGluIleAlaSerGly 317
|||||
|||||
901 GAATTACTCGATAATATGATTTATCAAAATTTGATTGAAATTCGATCTGG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
|||||
951 CGGAGCACCTTATCTAAGAAATGTGTGAAGCTGTGTGACGCTTTTA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAla 350
|||||
1001 ATTTACCGGTGTGTCGAAAGCTATGTTTAAACAGAAACACCTCTGCA 1050
351 lLeIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVal 367
|||||
1051 ATTATTATCACACCGGAGGCGATGATTAACCAAGGTCTTCTGCAAACT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
|||||
1101 TGTGCATTAATTTAAAGCAAAAGTTATCATCTGTGATCTAAAAAAAAT 1150
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
|||||
1151 TGGGCCCGAAGACAGCTGAGAAATTTGTGTAAGGCTCTATAGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
|||||
1201 AAAGTTATGTAGATATATCCAGAACAGACAGACAGAAATCATGATGAGA 1250
417 uGlyTyrPheHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
|||||
1251 AGGTTGTGTCACACAGAGATATGGGTATTAAGATGAAGAAAAAAT 1300
434 hPheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
|||||
1301 TCTTATCTGTGATCGCTTGAAGTCTTATCAATTAACAAAGCATATCA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIleP 467
|||||
1351 GTACCACTGTGAATTAAGATCTGTTTGGCAACATCCAAATATATTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLeuProG 484
|||||
1401 TGATGCCGCGCTGCTGCGGTCAGATCCTTAAGTCTGTCGCTCCGG 1450
484 lValAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
|||||
1451 GAGCTGTGTGTACTGAAAAAGGAAATCATGATGAAAAAGAAAGTA 1500
501 MetAspTyrValAlaSerGluValSerAsnAlaLysArgLeuArgGly 517
|||||
1501 ATGGATTACGTTGCTAGTCAAGTTCAAATGCAAAACCTTGGCGTGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
|||||
1551 TGTCCGTTTGTGCGAGCAATCACTAAAGGTCTCAGTGTAAATTTGACG 1600
534 lLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
|||||
1601 GTAAAGCAATTAAGAAATTAAGTGAAGAAACCAAGTTGCTAAAGATG 1644

```

seq_name: /SID2/gcgdata/geneseq/geneseqn/NA1997.DAT:AA163268

seq_documentation_block:

ID AA163268 standard; DNA: 1704 bp.

AA163268;

20-MAY-1997 (first entry)

Firefly Luciferase conjugated with 23 aa peptide at 5' end.

Fusion protein; firefly; lucifera; luciferase; biotinylation;
 wild type; E.coli; bioluminescence assay; quantification; ligand;
 receptor; ds.

XX

```

OS      Synthetic.
XX      Key
FH      CDS
FT      Location/Qualifiers
FT      1..75
FT      /tag= a
FT      /note= "encodes biotinylated peptide"
FT      CDS
FT      76..1704
FT      /tag= b
FT      /note= "encodes firefly luciferase residues 4-54"
XX      JP08308578-A.
XX      PD
XX      26-NOV-1996.
XX      PF
XX      24-APR-1995; 95JP-0098857.
XX      PR
XX      14-MAR-1995; 95JP-0054625.
XX      PR
XX      27-JUL-1994; 94JP-0193798.
XX      PA
XX      (KIKK ) KIKKOMAN CORP.
XX      DR
XX      WPI: 1997-059697/06.
XX      P-PSDB: AAM12394.
XX      PT
XX      Fusion protein comprising firefly luciferase and biotinylated
XX      peptide - useful in a bio-luminescent analytical method for
XX      quantifying ligands
XX      PS
XX      Example 1; Page 9-10; 13pp; Japanese.
XX      CC
XX      This is the nucleotide sequence encoding a novel fusion protein which
XX      comprises the firefly (Luciola lateralis) luciferase protein and a
XX      23 amino acid biotinylated peptide designated biotin peptide #84.
XX      CC
XX      The gene was generated by firstly annealing the sequence encoding
XX      peptide #84 (AAT63270) and its complementary sequence, then inserting the
XX      resultant double stranded sequence into the 5' end of the luciferase gene
XX      in plasmid pHLf108. The biotin peptide sequence replaces the first 4
XX      amino acids of the wild type luciferase sequence. The plasmid was then
XX      introduced into E.coli JM101 for production of the fusion protein. The
XX      novel protein can be used in bio-luminescence assays to quantify
XX      luciferase ligands which may modulate binding of luciferase to its
XX      receptor.
XX      SQ
XX      Sequence 1704 BP; 541 A; 275 C; 361 G; 527 T; 0 other:

alignment_scores:
      Quality: 2801.00      Length: 546
      Ratio: 5.130      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.451

alignment_block:
US-09-581-241-4 x AAT63268 ..

Align seg 1/1 to: AAT63268 from: 1 to: 1704

3 AsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrPr 19
: : : : : : : : : : : : : : : : : : : : : : : : : : :
67 AGTCTCGAAGCAATGAAATATGTCTATGCTGATCAACCATTTTATCCC 116
19 oileglugluglyseralaglyalaglnleuarglystymetaspargt 36
: : : : : : : : : : : : : : : : : : : : : : : : : : :
117 TATGAGAGGAGATCTGCTGACACAAATGCGCAAGTATATGATGATCAT 166
36 yralalysleuglyalalealapherhrasnaleuthrlyalasp 52
: : : : : : : : : : : : : : : : : : : : : : : : : : :
167 ATGCAAAACCTGGAGCAATGCTTTACTAACGACCTTACCGGCTGCAT 216
53 TyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGlyAlaLe 69
: : : : : : : : : : : : : : : : : : : : : : : : : : :
217 TATACGTACGCCGATATCTTATGAAATATCATGCTCTAGAGAGGCTTT 266
69 ulysasnTyrGlyLeuValValaspGlyArgIleAlaLeuCysSerGlu 86

```

```

267 AAAGATTTATGTTGTTGTTGATGAGAGAAATGCGTTATGACGTGAAA 316
86 snCysGluGluPhePheIleProValIleuAlaGlyLeuPheIleGlyAl 102
: : : : : : : : : : : : : : : : : : : : : : : : : : :
317 ACTGTGAAGAGTTCTTATTCCTGTATAGCCGGTTATATATATAGGGTGC 366
103 GlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHisSe 119
: : : : : : : : : : : : : : : : : : : : : : : : : : :
367 GGTGTGGCTCCACTAATGACATTTTACCTCTACCTGATGATGGTTCCACAG 416
119 rleuglylleserlyserprothrilevalpheserSerlysglyleua 136
: : : : : : : : : : : : : : : : : : : : : : : : : : :
417 TTTAGCATCTCTTAAGCCACAAATGATTTAGTTCTTAAAGAAAGATTAG 466
136 splysValIleThrValGlnLysThrValIThrAlaIleLysThrIleVal 152
: : : : : : : : : : : : : : : : : : : : : : : : : : :
467 ATAAAGTTATTAACCTGTACAAAAACGGTAACTGCTATTAAAAACCATGTT 516
153 ileuAspSerLysValAspTyrArgIlyTyrGlnSerMetAspAsnPh 169
: : : : : : : : : : : : : : : : : : : : : : : : : : :
517 ATATTGGACAGCAAGCTGATTTATAGAGTTATCATCATGACGACAACTT 566
169 eilleLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThrV 186
: : : : : : : : : : : : : : : : : : : : : : : : : : :
567 TATTAATAAAAAAACACTCCACAAAGTTTCAAGATCAAGATTAAAACTG 616
186 aIGluValAsnArgLysGluGlnValAlaLeuIleMetAspSerSerGly 202
: : : : : : : : : : : : : : : : : : : : : : : : : : :
617 TAGAAGTTAACCGCAAGAAAGAAAGTTGCTTATATATCACTTCCTGGGT 666
203 SerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLeuValTh 219
: : : : : : : : : : : : : : : : : : : : : : : : : : :
667 TCAACCGGTTTCCAAAGAGTGTGCACTTACTCATGAAATTTGGTCTAC 716
219 rArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValSerProG 236
: : : : : : : : : : : : : : : : : : : : : : : : : : :
717 GCGTTTTCCTCAGCTAGAGATCAATTTATGAAACCAAGTTTCACAG 766
236 lYThrAlaIleLeuThrValValProPheHisHisGlyPheGlyMetPhe 252
: : : : : : : : : : : : : : : : : : : : : : : : : : :
767 GCACGGCTATTTTAACCTGATGACATTCATCATGCTTTTGGTATGTTT 816
253 ThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLeuThrLy 269
: : : : : : : : : : : : : : : : : : : : : : : : : : :
817 ACTACTTAGGCTATCTACTGCTGTTTGGTATTCATGATTAACGAA 866
269 sPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSerS 286
: : : : : : : : : : : : : : : : : : : : : : : : : : :
867 ATTTGACGAAGAGACTTTTAAAAACACTGCAAGATTAACAATGTTCAA 916
286 erValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
: : : : : : : : : : : : : : : : : : : : : : : : : : :
917 GCCTTATCTCTGACGACTTGTGTTGCAATCTTATATGAAAGTGAATTA 966
303 LeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyAlaI 319
: : : : : : : : : : : : : : : : : : : : : : : : : : :
967 CTCGATTAATATGATTTATCAAAATTTAGTTGAATTTGCACTGGCGGAGC 1016
319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeuP 336
: : : : : : : : : : : : : : : : : : : : : : : : : : :
1017 ACCTTTATCTTAAAGAAATGTTGTAAGCTGTCTGACGCTTTTAAATTAC 1066
336 roglYValArgGlnGlyTyrGlyLeuThrGluThrThrSerAlaIleLe 352
: : : : : : : : : : : : : : : : : : : : : : : : : : :
1067 CGGGGTGTGCTCAAGGCTATGTTTACAGAAACAACCTCTGCATTTATTT 1116
353 lIeThrProGluGlyAspAspLysProGlyAlaSerGlyLysValValPr 369
: : : : : : : : : : : : : : : : : : : : : : : : : : :
1117 ATCACACCGGAAGGCGATGATTAACACAGGCTGCTTGGCAAAAGTTGTGC 1166
369 oleuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP 386
: : : : : : : : : : : : : : : : : : : : : : : : : : :

```



```

1167 ATTATTTAAAGCAAAAGTTATCGATCTGTACTATAAAAAAAGCTTGGGCC 1216
386  rAsnArgrArgIgluValCysValIysGlyProMetLeuMetIysGly 402
1217 CGAACGACGCTGAGAGAGTTTGTGTAAGGGTCTTATGCTTATGAAAGGT 1266
403  TyrValAspAsnProGluAlaThrArgGluIleIleAspGluGlyTr 419
1267 TATGTAGATAATCCAGAACCAAGAGAAATCATATGATGAGAAAGTTG 1316
419  pLeuHisThrGlyAspIleGlyTyrTyrTrpSpIuGluLysHisPhePhe 436
1317 GTTGACACAGGAGATATTGGTATTACGATGAGAGAAACATTTCTTTA 1366
436  leValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
1367 TCGTGATCGCTTGAAGTCTTTATCATCAATACAAAGATATCAAGTACCA 1416
453  ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAl 469
1417 CCTGCTGAATTAGAATCTGTTCTTTGCAACATCCAAATATTTTGATGC 1466
469  aGlyValAlaGlyValProAspProIleAlaGlyLeuLeuProGlyAlaV 486
1467 CGGCTTGCTGCGCTTCCAGATCTATAGCTGTAGCTTCGGGAGCTG 1516
486  alValAlaLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
1517 TTGTTGTACTTGAAGAAAGAAATCTATGACTGAAAGAAAGATATGAT 1566
503  TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValAr 519
1567 TACGTTGCTAGCAAGTTTCAATGCAAAACGTTGCGGTGCGTCCG 1616
519  pPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536
1617 TTTGTGACGACGATACCTAAAGTCTCAGTGTAAATGTGCGGTAAAG 1666
536  laIleArgGluIleLeuLysLysProValAlaLysMet 548
1667 CAATTAGAGAAATCTGAGAAACGATGCTAAGATG 1704

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV23595
seq_documentation_block:
ID  AAV23595 standard: cDNA to mRNA: 1704 BP.
XX
XX  AAV23595;
AC
XX
DT  16-JUL-1998 (first entry)
XX
DE  Antibody-Firefly luciferase fusion protein gene.
XX
KW  Firefly luciferase; antibody-luciferase fusion protein; ds.
XX
OS  Luciola cruciata.
XX
FH  key
FT  1..1704
FT  Location/Qualifiers
FT  CDS
FT  /tag= a
FT  /note= "no stop codon given"
XX
XX  JP09187281-A.
XX
XX  22-JUL-1997.
XX
XX  09-JAN-1996; 96JP-0001812.
XX
XX  09-JAN-1996; 96JP-0001812.
XX
XX  (KIRK ) KIRKMAN CORP.
XX
XX  PA
XX  WI: 1998-275089/25.
XX
XX  DR

```

```

DR  P-PSDB: AAW53883.
XX
XX  Antibody-firefly luciferase fused protein - and related products
PT  i.e. firefly luciferase fused gene, recombinant DNA and its
PT  preparation
XX
XX  Disclosure; Page 10-11; 17pp; Japanese.
XX
XX  This sequence encodes a fusion protein of the invention. The protein is
CC  a antibody-firefly luciferase fusion protein, in which an antibody part
CC  consisting of a peptide having antibody activity is combined with an
CC  enzyme part consisting of firefly luciferase.
XX
XX  Sequence 1704 BP; 541 A; 275 C; 361 G; 527 T; 0 other;
SO

alignment_scores:
Quality: 2801.00      Length: 546
Ratio: 5.130          Gaps: 0
Percent Similarity: 100.000      Percent Identity: 99.451

alignment_block:
US-09-581-241-4 x AAV23595      ..
Align seg 1/1 to: AAV23595 from: 1 to: 1704

3  AsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrPr 19
67  AGCTCGAAGACGATGAAATATGTGTAGTGTCTGAGCAATTTTACC 116
19  oIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetAspArgT 36
117  TATTAAGAGGAGGATCTGCTGAGCACAATTTGGCAAGTATATGATCGAT 166
36  yAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValasp 52
167  ATGCAAACTTGGACCAATGTCTTACTTAACGCATTAACCGGTCTGCAT 216
53  TyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyGluAlaIle 69
217  TATAGCTACGCCGATFACTTGAATAAATCATGCTCTTAAGAGAGGCTTT 266
69  uLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCysSerGluA 86
267  AAAGAAATATGTGTTGTTGTGATGAGAAATTCGTTATGCACTGAAA 316
86  snCysGluGluPhePheIleProValIleuAlaGlyLeuPheIleGlyVal 102
317  ACTGGAAGAGTTCTTATTCCTGTATTAGCCGCTTTATTTATAGTGTCTC 366
103  GlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHisSe 119
367  GGTGTGCTCCACTAATGAGATTTACACTCTACTGAAATTTGTTCAACG 416
119  rLeuGlyLysSerLysProThrIleValPheSerSerLysGlyLeuA 136
417  TTAGGCAATCTCTAAGCCACAATTTGATTAGGTTTAAAAAGGATTAG 466
136  sPlyValIleThrValGlnLysThrValThrAlaIleLysThrIleVal 152
467  ATTAAGTTATTAACGTGTACAAAAAACGGTAACTGCTATTAAAAACCAT 516
153  IleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAspAsnPh 169
517  ATATTGACAGCAAGTGATTTATAGAGTTATCATCATCATGACCACTT 566
169  eIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThrV 186
567  TATTAAAAAAACACTCCACAAGGTTTCAAGGATCAAGTTTAAAACTG 616
186  alGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSerSerGly 202
617  TAGAAGTTAACCCGCAAGACAAAGTGTCTTATTAATGAACCTTTCGGGT 666

```

```

203 SerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLeuValTh 219
|||||
667 TCACCGGTTTCCAAAGAGTGTGCAACTTACTCATGAAATTTGGTCC 716
219 rArgPheSerHisAlaArgAspProIleTyrGlyAsnGluValSerPro 236
|||||
717 GCCTTTTTCACGCTAGAGATCCAAATTTATGGAACCAAGTTTCACCG 766
236 lYThAlaIleLeuThrValValProPheHisHisGlyPheGlyMetPhe 252
|||||
767 GCACGGCCTATTTAACTAGTACATCCATCCATGCTTTGTGATGTTT 816
253 ThrThLeuGlyTyrIleuThrCysGlyPheArgIleValMetLeuThrly 269
|||||
817 ACTACTTTAGGCTATCTACTGTGGTTTTCGATTGTGATGTAACGAA 866
269 sPheAspGluGluThrPheLeuLysThrLeuGluAspTyrLysCysSer 286
|||||
867 ATTTCACGAAGAGACTTTTAAAACTGCACAGTTTACAAATGTTCA 916
286 eValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
|||||
917 GCGTATCTCTTGACGACTTTGTCATCTTATAGAGTGAATTA 966
303 LeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyLys 319
|||||
967 CTGCATTAATATGATTTATCAAAATTTAGTTGAATTCATCTGCGGAG 1016
319 aProLeuSerLysGluIleGlyLysAlaValAlaArgArgPheAsnLeu 336
|||||
1017 ACCTTATATCAAAAGAAATGTTGACAGCTGTGTCAGCTTTAATTTAC 1066
336 rGclYValAlaArgGlnGlyTyrGlyLeuThrGluThrThrSerAlaIle 352
|||||
1067 CGGGTGTCTTCACAGGCTATGTTTAAACAGAAACACCTCTGCAATTA 1116
353 lLethrProGluGlyAspAspLysProGlyAlaSerGlyLysValValPr 369
|||||
1117 ATCACACCGGAGAGCGCATGATTAACCGAGTCTTCGCAAACTTGCGCC 1166
369 OLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGly 386
|||||
1167 ATTATTTAAAGCAAAAGTATTCGATCTTGTACTAAAAAATTTGGGCC 1216
386 rAsnAlaArgGlyGlyValCysValLysGlyProMetLeuMetLysGly 402
|||||
1217 CGAACACAGCGGAGAGTGTGTTAAAGGTCCTATGCTTATGAAAGGT 1266
403 TyrValAspAsnProGluAlaThrArgGluIleLeuAspGluGlyTr 419
|||||
1267 TATGTAGTATATCCAGAACACAGCAAGAAATCTATGATGAAGAGTTG 1316
419 pLeuHisThrGlyAspIleGlyTyrTyrAspGluGlyLysHisPheMet 436
|||||
1317 GTTCACACAGAGAGATATGCGGTATGAGTGAAGAAAAAATTTCTTTA 1366
436 lValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
|||||
1367 TCGTGAGATCGTTTGAAGCTTTTAAATCAAAATPACAAAGATATCAAG 1416
453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAl 469
|||||
1417 CCTGTGATATAGAAATCTGTTCTTTGCAACATCCAAATATTTTGTATGC 1466
469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyLys 486
|||||
1467 CGGGCTTGTCTGCGTCCAGATCTATAGCTGTGAGCTTCCGGAGACTG 1516
486 aValValLeuLysGlyLysSerMetThrGluLysGluValMetAsp 502
|||||
1517 TTGTTGTTACTGAAAGAAATCTATGACTGAAAAAGAAAGTAAATGAT 1566

```

```

503 TyrValAlaSerGluValSerAsnAlaLysArgLeuArgGlyValAla 519
|||||
1567 TACGTGTCTACTCAAGTTTCAATGCAAAAGCTTTGGTGGTGGTCCG 1616
519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLys 536
|||||
1617 TTTTGTGACAGAGTACTTAAGCTTCACCTGTTAAATTCACGCTAAG 1666
536 lAlleArgGluIleLeuLysTyrProValAlaLysMet 548
|||||
1667 CAATTAGAAATCTGAAAGAACCAAGTTGCTAAGATG 1704
seq_name: /SID52/gcgdata/geneseq/geneseqn/NA1996.DAT:AA133850
seq_documentation_block:
ID AA133850 standard; DNA: 2019 BP.
XX
AC AA133850:
XX
DT 12-NOV-1996 (first entry)
XX
DE DNA encoding streptavidin/luciferase fusion protein.
XX
KW Streptavidin; luciferase; fusion protein;
KW Streptomyces avidinii; luciola lateralis; firefly;
KW recombinant production; industry; ds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..2019
FT /*tag= a
XX
PN JP07289264-A.
XX
PD 07-NOV-1995.
XX
PE 27-APR-1994; 94JP-0090275.
XX
PR 27-APR-1994; 94JP-0090275.
XX
PA (KIRK ) KIRKMAN CORP.
XX
WP1: 1996-015269/02.
DR P-PSDB: AAM04208.
XX
PT New mutant streptavidin (SA) gene - fused to a firefly luciferase
PT gene, for the recombinant prepn. of a SA-FL fused protein
XX
PS Example 1: Pages 5-6; 12pp; Japanese.
XX
CC The present sequence encodes a streptavidin/luciferase fusion
CC protein, comprising the Streptomyces avidinii streptavidin gene
CC and the luciola lateralis (firefly) luciferase gene. The fusion
CC protein can be prepd. by inserting the recombinant DNA encoding
CC it into a Escherichia species microbe, culturing the transformed
CC microbe in a medium and collecting the fusion protein from the
CC culture. The fusion protein can be used in industry.
XX
SQ Sequence 2019 BP: 590 A; 412 C; 467 G; 550 T; 0 other:

```

alignment_scores:

```

Quality: 2801.00 Length: 546
Ratio: 5.130 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.451

```

alignment_block:

```

US-09-581-241-4 x AA133850 ..

```

```

Align seg 1/1 to: AA133850 from: 1 to: 2019

```

```

.....
382 TCCCTCGAAGAGATGAAATATGTGTATGCTCTGAACCATTTTACC 431
19 011eGIuGIuGIySerAlaGIyAlaGlnLeuAryLysTyrMetAspArgT 36
432 TATTAAAGAGGATCTGCTGGAGCAATTCGCAAGATATATGATGCAT 481
36 yralAlaLysLeuGIyAlaIleAlaPheThrAsnAlaLeuThrGIyValAsp 52
482 ATGCAAACTTGGAGCAATTCGCTTTACTTAACGCACCTTACCGGTCTCAT 531
53 TyrThrTyrAlaGIuTyrLeuGIuLysSerCysCysLeuGIyAlaIle 69
532 TATACGTACGCCGATTAAGTAAAAATCATGCTCTAGAGAGAGCTTT 581
69 uLysAsnTyrGIyLeuValValAspGIyArgIleAlaLeuCysSerGIuA 86
582 AAAGATATATGTTGGTTGGTTGATGAGAGATTCGCTTATAGCAGTAAA 631
86 snCysGIuGIuPhePheIleProValLeuAlaGIyLeuPheIleGIyVal 102
632 ACTGTGAAGAGTCTTATTCCTGTATATAGCGCGTTATATTATAGGTGTC 681
103 GIyValAlaProThrAsnGIuIleTyrThrLeuArgGIuLeuValHisSe 119
682 GGTGTGCTCCAACTTAATGAGATTTACCTCTACGTAATGTGTCACAG 731
119 rLeuGIyLLeSerLysProThrIleValPheSerSerLysGIyLeuA 136
732 TTTAGGCACTCTTAAGCCAAATTTGATTAGTTCAAAAAAGATTAAG 781
136 sPlyValIleThrValGlnLysThrValThrAlaIleLysThrIleVal 152
782 ATTAAGTATTAACGTGTACAAAAAGGTAACCTGCTATTAAAAACATGTT 831
153 IleuAspSerLysValAspTyrArgGIyTyrGlnSerMetAspAsnPh 169
832 ATATTGACAGCAAGATGATTTAGAGTTATCAATCCATGCAGACATTT 881
169 eIleLysLysAsnThrProGlnGIyPheLysGIySerSerPheLysThrV 186
882 TATTAAAAAAAACACTCCACAGGTTTCAAAAGATCAAGTTTAAAAACTG 931
186 aGIuValAsnArgLysGIuGIuValAlaIleuIleMetAsnSerSerGIy 202
932 TAGAAGTTAAACCCGAAAGAAAGAGTGTCTTATTAATGAACCTTGCGGT 981
203 SerThrGIyLeuProLysGIyValGlnLeuThrHisGIuAsnLeuValTh 219
982 TCAACCGGTTGGCAAAAGGTGCAACTTACTCATGAATAATTTGGTCCAC 1031
219 rArgPheSerHisAlaArgAspProIleTyrGIyAsnGlnValSerProg 236
1032 GCGGTTTTTCACGCTAGAGATCCAAATTAATGAAACCAAGTTTCCACGAG 1081
236 lYThrAlaIleLeuThrValValProPheHisHisGIyPheGIyMetPhe 252
1082 GCACGGGTATTTAACTAGTAGTACATTCATCATGTGTTTGGTATGTTT 1131
253 ThrThrLeuGIyTyrLeuThrCysGIyPheArgIleValMetLeuThrLy 269
1132 ACTACTTATGCTATCTAACTTGTGTTTTCGATTTGCATGTATTAAGAA 1181
269 sPheAspGIuGIuThrPheLeuLysThrLeuGlnAspTyrLysCysSerS 286
1182 ATTTCAGCAAGAGACTTTTAAAAACACTGCACAGATTCAAAATGTTTCAA 1231
286 eValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGIuLeu 302
1232 GCGTATTTCTGTAGCGACTTTGTTGCAATTCCTTATATGAAGTGAATTA 1281
303 LeuAspLysTyrAspLeuSerAsnLeuValGlnIleAlaSerGIyGIyAl 319
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
1282 CTCGATTAATATGATTTATCAAAATTTAGTTGAATTCGACTCGGCGAGC 1331
319 aProLeuSerLysGIuIleGIyGIuValAlaValAlaArgArgPheAsnLeuP 336
1332 ACCTTATATCAAAAGAAATTTGTGACAGCTGTGTGCTACAGCTTTAATTTC 1381
336 rGIyValAlaArgGlnGIyTyrGIyLeuThrGIuThrThrSerAlaIleIle 352
1382 CGGGTGTGCTCAAGGCTATGTTTAAACAGAAACCAACCTCTCCAAATTAT 1431
353 lLeThrProGIuGIyAspAspLysProGIyAlaSerGIyLysValValPr 369
1432 ATACACACCGAAGCGATGATTAACACAGGTGCTTCGCAAACTGTGTC 1481
369 oLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGIyP 386
1482 ATTATTTAAAGCAAAAGTTATCATCTTCATCTAATAAAAAACCTTGCGCC 1531
386 roAsnArgArgGIyGIuValCysValLysGIyProMetLeuMetLysGIy 402
1532 CGACAGACAGCTGAGAGAGTTTGTGTTAAAGGCTCTATGCTTATGAAAGT 1581
403 TyrValAspAsnProGIuAlaThrArgGIuIleIleAspGIuGIyLys 419
1582 TATGTGATTAATCCAGAAACCAAGAGAAATCATAGATGAAGAAGTTG 1631
419 pleuHisThrGlyAspIleGIyTyrTyrAspGIuGIuLysHisPhePheI 436
1632 GTTGCAACAGAGAGATATGGGATTTACGATGAAAAAAACATTTCTTTA 1681
436 lValAspArgLeuLysSerLeuIleLysTyrLysGIyTyrGlnValPro 452
1682 TCGTGATGCTTTGAAGCTTTATCAATCAATCAAAAGATATCAATACCA 1731
453 proAlaGIuLeuGIuLysValLeuLeuGlnHisProAsnIlePheAspAl 469
1732 CCGTGGAATTAAGATCTGTTCTTTGCAACATCCAAATATTTTGTGATGC 1781
469 aGIyValAlaGIyValProAspProIleAlaGIyGIuLeuProGIyAlaV 486
1782 CGGCGTGTGCTGCTTCCAGATCCATATAGCTGTGTGAGCTTCCGGAGCTG 1831
486 aValValLeuLysLysGIyLysSerMetThrGlnLysGIuValMetAsp 502
1832 TTGTGTACTTGAAGAAAGAAATCTATGACTGAAAAAGAAAGATTAATGAT 1881
503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGIyGIyValAr 519
1882 TACGTGCTAGTCAAGTTTCAAAATGCAAAACGTTGCTGTGTGTGTCG 1931
519 gPheValAspGIuValProLysGIyLeuThrGIyLysIleAspGIyLysA 536
1932 TTTTGTGACGAAAGTAAAGTCTCACTGTAATAATTTACGCTAAAG 1981
536 lAlleArgGIuIleLeuLysLysProValAlaLysMet 548
1982 CAATTAGAAATTAAGTGAAGAAACAGTTGCTAAGATG 2019
seq_name: /SID2/gcgdata/geneseq/geneseqn/NA1996.DAT:AA133851
seq_documentation_block:
ID AA133851 standard; DNA; 2055 bp.
XX
AC AA133851;
XX
AT 12-NOV-1996 (first entry)
XX
DE DNA encoding mutant streptavidin/luciferase fusion protein.
XX
KW Mutant; streptavidin; luciferase; fusion protein;
KW Streptomyces avidinii; Luciola lateralis; firefly;
KW recombinant production; industry; ds.
XX

```

```

OS      Synthetic.
XX      Key      Location/Qualifiers
FH      CDS      1..2055
FT      /lag- a
FT      /note= "stop codon absent"
XX      JP07289264-A.
XX      07-NOV-1995.
XX      27-APR-1994; 94JP-0090275.
XX      27-APR-1994; 94JP-0090275.
XX      (KIKK ) KIKKOMAN CORP.
XX      WPI: 1996-015269/02.
XX      P-PSDB: AAM04208.
XX      New mutant streptavidin (SA) gene - fused to a firefly luciferase
XX      gene, for the recombinant prepn. of a SA-FL fused protein
XX      Claim 1: Pages 7-8; 12pp; Japanese.
XX      The present sequence encodes a mutant streptavidin/luciferase
XX      fusion protein, comprising a mutant Streptomyces avidinii
XX      streptavidin gene and the Luciola lateralis (firefly) luciferase
XX      gene. The fusion protein can be prepd. by inserting the
XX      recombinant DNA encoding it into a Escherichia species microbe,
XX      culturing the transformed microbe in a medium and collecting the
XX      fusion protein from the culture. The fusion protein can be used
XX      in industry.
XX      Sequence 2055 BP; 603 A; 418 C; 474 G; 560 T; 0 other:

alignment_scores:
      Quality: 2801.00      Length: 546
      Ratio: 5.130      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 99.451

alignment_block:
US-09-581-241-4 x AAT33851 ..
Align seg 1/1 to: AAT33851 from: 1 to: 2055

3 AsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrPr 19
  ::::::::::::::::::::::::::::::::::::::::::::::::::::
418 TCCCTCGAAGACGATGAAATATTTGCTATGTCGACGACCATTTACCC 467
19 cllleugluclySeraIaGlyAgluInleuArglyTyrMetAspArgT 36
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
468 TATTGAAGAGGATCTGCTGACACAAATGCGCAAGTATATATATTCGAT 517
36 yralaIySleuGlyAlaIleAlaPheThrAsnAlaIeuThrGlyValAsp 52
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
518 ATGCAAAACTTGAGCAATGCTTTTACTAACGCACTTACCGGTTCGAT 567
53 TyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGlyGluAla 69
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
568 TATACGTACGCCCGAATCTTAGAAAAATCATGCTGCTAGAGAGGCTTT 617
69 ulysAsnTyrGlyIleuValAlaIleAlaPheThrGlyValAsp 86
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
618 AAAAATTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 667
86 snCysGluGluPhePheIleProValIleuAlaGlyIleuPheIleGly 102
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
668 ACTGTGAAGATTCTTATTCCTGTTATAGCCGTTTATTTATAGTGTC 717
103 GlyValAlaProThrAsnGluIleTyrThrLeuArgGluIleuValHis 119
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::

```

```

718 GGTGTGCTCCAACTAATGAGATTTACACTCTACGTGATTTGGTTCACAG 767
119 rleuGlyIleSerLysProThrIleValIlePheSerLysGlyIleuA 136
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
768 TTTAGGATCTCTAAGCCAACTGATTTAGTTGTTGTTGTTGTTGTTGTT 817
136 sPlsValIleThrValGlnIleThrValThrAlaIleIleThrIleVal 152
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
818 ATTAAGTTATACCTGTACAAAAAAGCGTAACTGCTATTAACCATGCT 867
153 IleuAspSerLysValAspTyrArgGlyTyrGlnSerIleAspAsn 169
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
868 ATATTGACACCAAGTGATTTAGATTCATTCATTCATTCATTCATTCAT 917
169 eIlleLysAsnThrProGlnGlyPheLysGlySerSerPheLysThr 186
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
918 TATTAAAAAAACACTCCACAGGTTTCAAGAGTCAAGTTTAAAACTG 967
186 aGluValAsnArgLysGluIleuValAlaIleuIleMetLysSerGly 202
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
968 TAGAAGTTAACCCGCAAGACAGTGTGCTTATATGAACTCTTCGGGT 1017
203 SerThrGlyLeuProLysGlyValGlnIleuThrHisGluAsnIleu 219
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1018 TCACCGCTTTGCCAAAAGGTGCAACTCTCATGATAAAATTTGGTTCAC 1067
219 rArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValSerPro 236
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1068 GCGTTTCTCAGCGTACAGATCCATTTATGGAACCAACTTTCACACAG 1117
236 lYThrAlaIleuThrValAlaProPheHisHisGlyPheGlyMetPhe 252
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1118 GCACGCTATTTTACCTAGACATCCATCATGCTTTGGTATGTTT 1167
253 ThrThrIleuGlyTyrLeuThrCysGlyPheArgIleValIleMet 269
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1168 ACTCTTTAGGCTATCTAATCTGTGTTTGTATGTCATGATTAACGA 1217
269 sPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSer 286
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1218 ATTGACGACAGACTTTTAAACACTGCAAGATTAACAAGTTCGCA 1267
286 erValIleuValProThrIleuPheAlaIleuAsnArgSerGluIleu 302
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1268 GCGTATCTCTGTACCGACTTTGTTGCAATCTTAATGAAATCAATTA 1317
303 leuAspLysTyrAspLeuSerAsnIleuValGluIleAlaSerGlyVal 319
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1318 CTCGATTAATATGATTTTCAATTTAGTTGAATTCATCTGCGCGAGC 1367
319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeu 336
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1368 ACCTTATCTAAGAAATGTTGGAAGCTGTCTGACGTTTATTTAC 1417
336 roGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSeraIleIle 352
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1418 CCGGTGTTGTCAGAGCTATGTTTAAACAGAAACCTCTGCAATTTT 1467
353 IleThrProGluGlyAspAspLysProGlyAlaSerGlyLysValAla 369
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1468 ATCACCCGGAAGCGCATATTAACAGAGTCTTGTGCAAAAGTTGTGCC 1517
369 oleuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGly 386
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1518 ATTATTTAAGCAAAAGTTATCGATCTTATCTAATAAAACTTGGGCC 1567
386 roAsnArgArgGlyGluValLysValLysGlyPrometIleMetLysGly 402
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1568 CGAACAGAGCTGAGAGAGTTTGTAAAGGTCATGCTTATGAAAGT 1617
403 TyrValAspAsnProGluAlaThrArgGluIleIleAspGluGlyTyr 419
  ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1618 TATGTATATATCCAGAACCAACAGAAATCATGATGAAGAAAGTTTG 1667

```

```

419 pleuHisThrGlyAspIleGlyTyr-TyrAspGluGluLysHisPhePheI 436
1668 GTTGACACAGAGATATTGGATTACGATGAGAAAGAAACATTTCTTTA 1717
436 IValAspArgLeuLysSerIleLysTyrLysGlyTyrGlnValPro 452
1718 TCGTGATCGCTTGAAGTCTTTAACTCAATCAAAAGGATATCAAGTACCA 1767
453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAl 469
1768 CTGCTGATTAAGATCTGTTTTCACACATCCAAATATTTTGATGCG 1817
469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyAlaVal 486
1818 CGGCGTGTGCGCTCCAGATCCTATAGCTGAGCTCCGAGACTCGAGACTG 1867
486 AlValAlaValLeuLysGlyLysSerMetThrGluLysGluValMetAsp 502
1868 TTGTTGCTACTGAAAGAAAGAAATCTATGACTGAAAGAAAGATATGAT 1917
503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValAlaVal 519
1918 TACGTGTCTACTCAAGTTTCAATGCAAAACGTTGCGTGTGTGTCCG 1967
519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536
1968 TTTTGTGACAGAGTACCTTAAGGTCTACGTGTAATAATTGACGCGTAA 2017
536 IAlaIleArgIleLeuLysLysProValAlaLysMet 548
2018 CAATTAGAGAAATACGAGAAACCACTGTGCTAAGATG 2055
seq_name: /SDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV23580
seq_documentation_block:
ID AAV23580 standard; cDNA to mRNA; 2364 BP.
AC AAV23580;
XX
XX
DT 16-JUL-1998 (first entry)
XX
DE Antibody-Firefly Luciferase fusion protein gene.
XX
XX Firefly Luciferase; antibody-Luciferase fusion protein; ds.
XX
XX Luciola cruciata.
XX
XX
XX Key Location/Qualifiers
XX CDS 1..2364
XX FT /tag- a
XX FT /transl_except- (pos: 670..672, aa: Gln)
XX FT /transl_except- (pos: 739..741, aa: Trp)
XX FT /transl_except- (pos: 1369..1371, aa: Ala)
XX FT /note= "no stop codon given"
XX
XX JP09187281-A.
XX
XX PD 22-JUL-1997.
XX
XX PF 09-JAN-1996; 96JP-0001812.
XX
XX PR 09-JAN-1996; 96JP-0001812.
XX
XX PA (KIKK ) KIKKOWAN CORP.
XX
XX DR WPI: 1998-275089/25.
XX
XX DR P-PSDB; AAM53882.
XX
XX Antibody-firefly Luciferase fused protein - and related products
XX PT i.e. firefly Luciferase fused gene, recombinant DNA and its
XX PT preparation
XX

```

```

PS Disclosure: Page 13; 17pp; Japanese.
XX
XX This sequence encodes the fusion protein of the invention. The protein is
CC a antibody-firefly Luciferase fusion protein, in which an antibody part
CC consisting of a peptide having antibody activity is combined with an
CC enzyme part consisting of firefly Luciferase.
XX
XX Sequence 2364 BP; 697 A; 444 C; 559 G; 664 T; 0 other;
SQ
alignment_scores:
Ratio: 2800.00 Length: 545
Percent similarity: 100.000 Gaps: 0
Percent identity: 99.633
alignment_block:
US-09-581-241-4 x AAV23580 ..
Align seg 1/1 to: AAV23580 from: 1 to: 2364
4 MetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrProI 20
730 CTGAGAACGATGAAATTTGTGTATGCTCCTGAACCATTTTACCTAT 779
20 eGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetAspArgTyrA 37
780 TGAAGAGGATCTGCTGAGACAAATGGCGCAAGTATATGATGATGATG 829
37 IAlsLeuGluValAlaIleAlaPheThrAsnAlaLeuThrGlyValAspTyr 53
830 CAAACTTGAGCAATGCTTTACTTAACGCACTTACCGGTGTCCATTAT 879
54 ThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyAlaAlaLeuLys 70
880 ACGTACGCCGATACTTAGAAATATCATGCTGTACGAGAGAGCTTTAA 929
70 sAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuCysSerGluAsn 87
930 GAATTATGTTGTTGTTGTTGATGAGAAATGCGGTATGACATGAAACT 979
87 ySGLuGluPhePheIleProValIleuAlaGlyLeuPheIleGlyValGly 103
980 GTGAGAGTCTTATTCCTGATTAAGCGGTTTATTAAGGTGTCGCT 1029
104 ValAlaProThrAsnGluIleTyrThrLeuArgIleuValHisSerle 120
1030 GTGCTCCCACTAATGAGATTACCTACGTGATGTTGTTACAGATT 1079
120 uGlyIleSerLysProThrIleValPheSerSerLysLysGlyLeuAsp 137
1080 ACGCATCTTAAGCCACAAATGTTATGTTCTTAAGAAAGATTAGATA 1129
137 ySValIleThrValGlnLysThrValThrAlaIleLysThrIleValIle 153
1130 AAGTATATACGTGTACAAAAACGGTAACGTCTATTAAACCATTTGTA 1179
154 LeuAspSerLysValAspTyrArgIlyTyrGlnSerMetAspAsnPheI 170
1180 TTGGACAGCAAGGATTAAGATTATCAATCATGACCAACTTTAT 1229
170 eLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThrVal 187
1230 TAAAAAAAACACTCCACAAGTTTCAAGATCAAGTTTAAACCTGTAG 1279
187 IValAlaAsnArgLysGluGlnValAlaLeuIleMetAsnSerSerGlySer 203
1280 AAGTTAACCGCAAGAACAAAGTCTCTTATTAATGAACCTCTCGGGTTCA 1329
204 ThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLeuValThr 220
1330 ACCGGTTTGCCAAAAGGTGTCACACTTACATCAATTAATTTGGTCACGCG 1379
220 gPheSerHisAlaArgAspProIleTyrGlyAsnGlnValSerProGlyT 237

```



```

34  spargtyrAlaLysLeuGlYAlaIleAlaPheThrAsnAlaLeuThrGly 50
    |||||
101  ATGATATGCAAAACTTGAGCAATTCGCTTACTACCACTTACCGGT 150
    |||||
51  ValAspTyrThrTyrAlaGluTyrLeuGluYysSerCysLeuGlyYcI 67
    |||||
151  GTGATTTATAGCTAGCCGAATCTAGTAAATAATCATCTGCTCTGAGGA 200
    |||||
67  uAlaLeuYsAsnTyrGlyLeuValIAspGlyArgIleAlaLeuCysS 84
    |||||
201  GGCCTTAAGAATATATGCTTGGCTTGTGATGAGAGAAATGCGTTATGCA 250
    |||||
84  erGluAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
    |||||
251  GTGAAACTGTGAAGAATCTTTATTCCTGATATAGCCGGTTATTTATA 300
    |||||
101  GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVa 117
    |||||
301  GGTGCGGTGGCTCCCACTAATGAGATTACACTCTACGTAAGTAATGCT 350
    |||||
117  LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
    |||||
351  TCACAGTTTAGCAATCTCTAGCCACAACTTGTATTTAGTTCTAATAAAG 400
    |||||
134  LysLeuAspLysValIleThrValGluYsThrValThrAlaIleLysThr 150
    |||||
401  GATTAGATAAAGTTATTAAGTCTACAAAACGCTAATCTATTAAAC 450
    |||||
151  IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
    |||||
451  ATTGTTATTTAGACAGCAAGAGATTAAGAGTTATACATTCATGCA 500
    |||||
167  PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
    |||||
501  CAACCTATTATTAATAAAACACTCCACAAGTTTCAAGGATCAAGTTTAA 550
    |||||
184  ystThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
    |||||
551  AAACGTGAGAAAGTTAACCGCAAAAGACAAAGTTGCTTATTAATCAAC 600
    |||||
201  SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnIle 217
    |||||
601  TCGGGTTCAACCGCTTGCCAAAAGGTGTGCACTTACTCATGAAATTT 650
    |||||
217  uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
    |||||
651  GGTCACTAGATTTTCTCAGCTAGAGATCCAAATTTATGGAACCAAGTTT 700
    |||||
234  erProGlyThrAlaIleLeuThrValAlaProPheHisHisGlyPheGly 250
    |||||
701  CACGAGCGCGCTATTTTAACGTAGTACCATTCATCATGTTTGGT 750
    |||||
251  MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIle 267
    |||||
751  ATGTTACTACTTATAGCTATCTAAGTGTGCTTGTGATTTGTCATGTT 800
    |||||
267  uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
    |||||
801  AACGAAATTTTACGCAAGAGACTTTTAAAAACACTCCAAAGATTACAAAT 850
    |||||
284  ysserSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
    |||||
851  GTTCAMCGCTTATTTCTGTACCGACTTGTGTAATCTTAAAGAGT 900
    |||||
301  GluLeuLeuAspLysTyrAspLysSerAsnLeuValGluIleAlaSerGly 317
    |||||
901  GAATTAATCATGAATAATGATTTATCAAAATTTAGTTGAATTTGATCTGG 950
    |||||
317  yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
    |||||
951  CGGAGCACCTTTATCTAAAGAAATGGTGAAGCTGTGTGAGACGTTTAA 1000
    |||||

```

```

334  snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
    |||||
1001  ATTTACGGGGTGTGCTCAGGCTATGTTTAAAGAAACCACTCTGCA 1050
    |||||
351  IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
    |||||
1051  ATTATTAACACACCGGAAGCGCATGATAAACAGAGTGTCTCGCAAACT 1100
    |||||
367  lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
    |||||
1101  TGTGCTATTATTTAAAGCAAAAGTTATGATCTTATCTAATAAAAACTT 1150
    |||||
384  euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
    |||||
1151  TGGCGCCCAACAGACGTGAGAGAGTTTGTGTAAGGCTCTATGCTTATG 1200
    |||||
401  LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluG 417
    |||||
1201  AAAGTTATGTAGATAATCCAGAACCAAGAGAAATCATAGATGAGA 1250
    |||||
417  uGlyThrPheuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
    |||||
1251  AGTTGTGTGACACAGAGATATTGGGTATTACGATGAAGAAAAACATT 1300
    |||||
434  hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
    |||||
1301  TCTTATGCTGATGCTGTTGAAGTCTTTATCAATACAAAGATATATAG 1350
    |||||
451  ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
    |||||
1351  GTGGCCCCGCTGAATTGGAATTCATATTGTACAAACCCCAACATCTT 1400
    |||||
467  eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
    |||||
1401  CGAGCGGGCGCTGCGAGCTTCTCCGACGATGACGCGGTAACTCCG 1450
    |||||
484  lValAlaValAlaValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
    |||||
1451  CGGCGGTGTTGTTTGGAGCAGCAAGAGATGACGAAAGAAAGAGATC 1500
    |||||
501  MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyYcI 517
    |||||
1501  GTGATTAACGTGCGCAGTCAAGTAACAACCGCAAAAGTTGCCGAGG 1550
    |||||
517  yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
    |||||
1551  AGTTGTGTTGTGACGAGAAAGTACCGAAAGTCTTACCGGAAACTCGACG 1600
    |||||
534  lLysAlaIleArgGluLeuLeuLysLys 543
    |||||
1601  CAAAGAAATTCAGAGATCTCTAATAAG 1629
    |||||
seq_name: /SID2/gcgsdata/geneseq/geneseqn/AA1989.DAT:AA91170
seq_documentation_block:
ID  AA91170 standard; DNA; 1644 BP.
AC  AA91170;
XX
XX  14-JUN-1990 (first entry)
DE  Recombinant luciferase gene.
XX
XX  Luciferase; ATP assay; ss.
XX
XX  Luciola cruciata.
OS
XX  EP301541-A.
XX
XX  01-FEB-1989.
XX
XX  28-JUL-1988; 88EP-0112233.
XX

```


PR 29-JUL-1987; 87JP-0187724.
PR 29-JUL-1987; 87JP-0187725.
PR 20-AUG-1987; 87JP-0205194.

(KIKK) KIKKOMAN CORP.

PI Masuda T, Tatsumi H, Nakano E;

DR WPI: 1989-033443/05.
P-PSDB: AAP94367.

XX Recombinant luciferase gene - used for efficiently producing luciferase
in E. coli, obtaining a stable and highly active prod.

PS Claim 2; fig 3; 35pp; English.

CC The sequence was obt'd. from cDNA prepd. from mRNA isolated from
L. cricetata tails. The DNA can be incorporated into vectors for
transformation of E. coli JM 1010.

CC Sequence 1644 BP; 529 A; 276 C; 338 G; 501 T; 0 other;

alignment_scores:

Quality: 2684.00 Length: 548
Ratio: 4.943 Gaps: 0
Percent Similarity: 99.088 Percent Identity: 93.431

alignment_block:

US-09-581-241-4 x AAN91170

Align seg 1/1 to: AAN91170 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProH 17
1 ATGGAAACATGGAAACGATGAAATATGTGAGTGGACCTAACCGTT 50
17 eTyrProIleGluGlySerIleGlyAlaGluIleuArgLysTyrMet 34
51 TTACCCATTCGAAAGAGGATCTCTGACACACATTCGCAATACATG 100
34 sPArGlyAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 AGCGATTCGAAACATGGCGCAATGCTTTACAATGCAATGCTACTGG 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
151 GTGATATATCTTACCGCGAATCTTGAGAAATCATGTTGCTAGGAAA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
201 AGCTTTCGAAATATGAGTTGGTTGGTGGACGAAATGCGGTATGCA 250
84 eArgLysCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAACTGTGAGAAATTTTATCTCTGTAATAGCCGCACTGTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleGlyThrLeuArgGluLeu 117
301 GGTTAGGTGTTGCACCCACTAATGAGATTTACCTTACGTAACCTG 350
117 IHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLys 134
351 TCACAGTTTAGGTATCTCTAACCACACATGTAATTTAGTTCTAAAA 400
134 ILeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
401 GCTTAGATAAGTATATACAGTACAGAAACAGTAACATATATAAAC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
451 ATTGTATATCTAGATGACAAAGTGTATATGAGGATATCAATGTCTGA 500

167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
501 CACCTTTATATAAAGAAACACTCCACACAGTTTCAAGATCCAGTTTCA 550
184 ySThrValGluValAsnArgLysGluGluValAlaLeuIleMetAsnSer 200
551 AAAGTGTGAAAGTTGACCGTAAGAACAAAGTCTGCTTATATGAACCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
601 TCGGTTCTACCGGTTTGGCAAAAGGCGTACAACTTACTACAGAAATAC 650
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
651 AGTCACATGATTTCTCATGCTAGAGATCCGATTTATGTAACCAAGTT 700
234 eProGlyThrAlaIleLeuThrValValProPheHisGlyPheGly 250
701 CACGAGGACCGCTTTTACCTGCTCCATTCATCCATGATGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMet 267
751 ATGTCACCTACTAGCGTATTTATTTGCTTTGCTGCTGTAATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLys 284
801 AACAAATTCGATGAGAAACATTTTAAAACTTACAAAGATATTAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTACAAAGCTTATCTTGTACCGACCTTGTGCAATTCATCAACAAAGT 900
301 GluLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
901 GAATTAATCAATTAATGATTTGCAATTTAGTGAATTCGATCGCG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
951 CCGAGCACCTTATCAAAAGAAAGTGTGAAAGCTGCTGACGCTTTA 1000
334 sLeuProGlyValArgGlnGlyTyrGlyLeuThrGlnThrSerAla 350
1001 ATCTTCGCGGTGCTGCAAGGTTATGTTTACGAAACAAACATCTGCC 1050
351 IleIleIleThrProGluLysAspLysProGlyAlaSerGlyLysVal 367
1051 ATTATATATACACAGAAAGACAGATTAACAGAGCTTCGCAAAAGT 1100
367 IValProLeuPheLysAlaLysValIleAspLeuSerThrLysThr 384
1101 CGTCCGCTGTTTAAAGCAAAAGTATGATCTTCAATACCAAAATCTT 1150
384 eGlyProAsnArgArgGlyGluValLysValLysGlyProMetLeuMet 400
1151 TAGGTCTTAACAGACGTGAGAAAGTTTGTAAAGACCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGln 417
1201 AAAGTTATGTAATTAATCCAGAACACAAAGAACTTATGACGAAACA 1250
417 uGlyThrLeuHisThrGlyAspIleGlyTyrGlyArgGluLysHisP 434
1251 AGGTGGCTGCACACCGAGATATTTGGATATTTATGTAAGAAAAACAT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTTATGTCGATCGTTGAAGCTTTTATCAATATCAAAAGATACCAA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIleP 467
1351 GTACACACCTCCGAATTAGAATCCGTTCTTTGCAACATCCATCTATCT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuPro 484

```

|||||
1401 TATGCTGTTGTCGCCGGCTTCGATCCTGCTAGCTGCGAGCTTCAG 1450
484 1YAlaValValleuValysGlyLysSerMetThrGluysGluVal 500
1451 GAGCGGTTGTTGCTACTGAAACGGAAGAAAAATATGACGAAAGAACTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
1501 ATGGATTATGTTGCAAGTCAAGTTTCAATGCAAAACGTTTACGTGTG 1550
517 YValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
1551 TGTTCGTTTGTGATGAGTACCTTAAGAGTCTTACGGAATAATGACG 1600
534 1YValAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
1601 GCAGAGCAATTAGAGAAATCCTTAAGAAACAGTTGCTAGATG 1644
seq_name: /SDS2/gcgdata/geneseq/geneseq/NA1990.DAT.AA03801

```

seq_documentation_block:

ID AA03801 standard; cDNA; 1644 BP.

AC AA03801;

DT 22-AUG-1990 (first entry)
DT 26-FEB-1993 (revised entry)

DE Recombinant luciferase gene.

KW Luciferase; ATP; E. coli; photon; ds.

OS Synthetic.

PN JP02065780-A.

PD 06-MAR-1990.

PF 01-SEP-1988; 88bp-0216229.

PR 01-SEP-1988; 88bp-0216229.

PA (KIKK) KIRKMAN CORP.

DR WPI; 1990-113360/15.

DR P-PSDB; AAR05788.

XX Prepn. of luciferase -
XX using Escherichia sp. bacteria contg. recombinant DNA.

XX Claim 2; Page 416 + Fig 3; 20pp; Japanese.

XX Prepn. of luciferase comprises culturing Escherichia sp. contg.
XX recombinant DNA, and harvesting luciferase from the culture medium.
XX The recombinant DNA is a vector DNA into which has been inserted the
XX luciferase gene which has the nucleotide sequence below. Luciferase
XX can be produced quickly and efficiently with an increased photon
XX output, useful in assaying ATP.

SO Sequence 1644 BP; 529 A; 276 C; 338 G; 501 T; 0 other;

alignment_scores:

Quality: 2684.00 Length: 548
Ratio: 4.943 Gaps: 0
Percent Similarity: 99.088 Percent Identity: 93.431

alignment_block:

US-09-581-241-4 x AA03801 ..

Align seg 1/1 to: AA03801 from: 1 to: 1644

```

1 MetLysMetGluAsnAspGluAsnIleValTyrGlyProGluProph 17
1 ATGGAAACATGAGAAACGATGAAATATGTGATGAGCTTAACCGTT 50
17 eTyrProIleGluGlySerAlaGlyAlaGlnLeuAspGlyTyrMet 34
51 TTACCCATGGAAGAGGATCTCTGGACACATTAACCAATTAATG 100
34 sPArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 AGCGATATGCAAAATCTGGCGCAATTGCTTTACAATCCAGTTACTGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGlyLysSerCysGlyLeuGly 67
151 GTTGATATCTTACCGCCGAAATCTTGGAATACTGTTGCTAGGAAA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuGly 84
201 AGCTTGCAAAATATGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTG 250
84 eTGLuAsnCysGluGluPhePheIleProValIleuAlaGlyLeuPheIle 100
251 GTTGAACACGTGAGAAATTTTATTCCTGTAATAGCCGACTGTTTAA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuAspGluLeuVal 117
301 GGGTAGAGTGTGTCACCCACTAATGAGATTACCTTACGTGAACCTGT 350
117 HisSerLeuGlyIleSerLysProThrIleValPheSerLysGly 134
351 TCACAGTTAGTATCTCTAAACCAACAATGTTATTTAGTTCTAAAG 400
134 1YLeuAspLysValIleThrValGlnLysThrValAlaIleLysThr 150
401 GCTTAGATTAAGTTATTAACAGTACGAAACAGTACTACTTAAAC 450
151 1YValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
451 ATTGTTATCTAGTATGCAAGATTGATATGAGGATATCAATGCTGCA 500
167 pAsnPheIleLysLysAsnThrProGlnGlyPheLysSerSerPheIle 184
501 CACCTTTATTAAGAAACACTCCACAGGTTTCAACGATCCAGTTTCA 550
184 ySThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAACGTGGAAGTTGACCGCTAAGACAGTTGCTTATATATCACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLe 217
601 TCGGTTCTACCGGTTTCCAAAAGCGGTACACCTACTCAGAAATATAC 650
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
651 AGTCACTAGATTTCTCATGCTAGACATCCGATTTATGGTAAACCACTT 700
234 ePProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
701 CACGAGGACCGCTGTTTAACTGCTGCATTCATCATAGGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
751 ATGTTCACTACTAGGGTATTTAATTTGGGTTTTCGTGTTAAATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
801 AACCAAAATTCGATGAAGAAACATTTTAAAACTCTACAAAGATTATTAAT 850
284 ySSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTACAACTGTTATTTCTGACGACCTTGTTCATTTGCAATTCGCAACAAAG 900
301 GluLeuAspLysTyrIleAspLeuSerAsnLeuValGluIleAlaSerG 317

```

```
|||||:|||||901 GAATTAAGTAAATACATTTGTCAAATTTAGTGAGATTCGACATCTGC 950
|||||:|||||317 YGlyAlaProLeuSerLysGluIleGlyAlaValAlaArgArgPheA 334
|||||:|||||951 CGAGACACCTTATCAAAAGAGTTGTGAAGCTTTGCTAGACGCTTAA 1000
|||||:|||||334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
|||||:|||||1001 ATCTTCCCGTGTGTCAGAGTTATGTTTACAGAAACAAACATCTGCC 1050
|||||:|||||351 llelleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVal 367
|||||:|||||1051 ATTATTATTACACAGAGAGAGCATTAACCGAGAGCTTCGAGAAAGT 1100
|||||:|||||367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
|||||:|||||1101 CGTGCCGTTGTTAAAGCAAAAGTTATGATCTTGATACCAAAAATCTT 1150
|||||:|||||384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
|||||:|||||1151 TAGGTCCTACAGACGTCGAGAGTTTGTGTTAAAGACCTATGCTTATG 1200
|||||:|||||401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluI 417
|||||:|||||1201 AAAGTTATGTAATATATCCAGAACACAAAGAACTTATTGACGACAG 1250
|||||:|||||417 uGlyTyrPheuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
|||||:|||||1251 AGGTGGCTGCACACCGAGATTTGATATGATGAAGAAACATTT 1300
|||||:|||||434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
|||||:|||||1301 TCTTTATGTGAGATCGTTGAGATCTTTATCAAAATACAAAGATACCA 1350
|||||:|||||451 ValProProAlaGluLeuGluSerValIleuLeuGlnHisProAsnIlePh 467
|||||:|||||1351 GTACCACTGCGCAATTAGAATCCGTTCTTTCACACATCCATCTATCTT 1400
|||||:|||||467 eaSPAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
|||||:|||||1401 TGATGCTGCTGTGCCGCGCTTCCTGATCCTGTAGCTGGGAGCTTCAG 1450
|||||:|||||484 lValAlaValIleuLysLysGlyLysSerMetThrGluLysGluVal 500
|||||:|||||1451 GAGCCGTTGTTGTAAGAGCGGAAAAATATGACCGGAAAAAGAAATG 1500
|||||:|||||501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
|||||:|||||1501 ATGATTTATGTTGCAAGTCAAGTTTCAAAATGCAAAACGTTTACGTG 1550
|||||:|||||517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
|||||:|||||1551 TGTTCGTTTGTGTGATGAAGTACCTTAAGGTCTTACTGGAATAATTGAC 1600
|||||:|||||534 lLysAlaIleArgGluIleuLysLysProValAlaLysMet 548
|||||:|||||1601 GCAGAGCAATTAGAGAAATCTTAGAAACCAAGTTCCTAAGATG 1644
```


34 spArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
 101 ATCGATATGCAAAACTGGACCAATGCTTTACTAAGCACTTACCGGT 150
 51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyGly 67
 151 GTGATATATAGTACGGCAATGCTTACGAAAAATCATGCTGTTAGGAGA 200
 67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
 201 GGCTTTAAAGAAATATAGTGTGGTTGTTGATGAGAAATTCGTTATGCA 250
 84 ergLysAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
 251 GTGAAACTGGAAGAATCTTTATCTCTGATATAGCCGGTTATTTATA 300
 101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
 301 GGATCGGTGGTCCCACTCAATGAGATTTACCTACGTCGAATGGT 350
 117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLysG 134
 351 TCACAGTTTGGGCACTCTTAAGCCACAATTTGATTTAGTTCTAAAAAG 400
 134 LysLeuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150
 401 GATTAGTAAAGTTATTAACCTACAAAAAAGCGTAACCTGCTATTTAAACC 450
 151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetLys 167
 451 ATGTGTTATTTGGACACCAAGAGATTATGAGGTTATCATCAATCCATGGA 500
 167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheIle 184
 501 CACTTTATTTAAAAAACAACCTCCACAGGTTCAAGATCAAGTTTAA 550
 184 YThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
 551 AAACGTGAGAGTTAACCCGAAAGAACAAAGTGTCTTATTAATGAACTCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnIle 217
 601 TCGGTTCAACCGGTTTGCACAAAAGGTGCACTTACTCATGAATAT 650
 217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
 651 CGTCACTAGATTTTCTACGCTAGAGATCCAAATTTATGAAACCAAGTTT 700
 234 ePProGlyThrAlaIleLeuThrValValProPheHisGlyPheGly 250
 701 CACGAGGACGCGCTATTTTACTGTAGTACCATTCATGCTGTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
 751 AAGTTTACTACTTTAGGCTATCTAATCTGTTGTTGTTGTTGTTGTTGTT 800
 267 unThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
 801 AACGAAATTTGACGAGAGACCTTTTAAAAACATGCAAGATTACAAAT 850
 284 YSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 851 GTTCAAGCGTTATCTTGTACCGACACTTTGTTGCAATTTCTTAATTAAGA 900
 301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGly 317
 901 GAATTAAGTGAATAATATGATTTATCAAAATTTAGTGAATTCATCCTGG 950
 317 yGlyValProLeuSerLysGluIleGlyValAlaValAlaArgArgPheA 334
 951 CGGAGCACCTTATCTTAAGAAATGTTGTAAGCTGTTGCTAGACGTTTAA 1000

334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAla 350
 1001 ATTTACCGGGGCTTCTGTTAAGGCTATGTTTACAGAAACCACTGCA 1050
 351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVal 367
 1051 ATTATTTATCACACCGGAAGCGGATGATTAACAGGTGCTTCTGCCAAGT 1100
 367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysIleThrL 384
 1101 TGTCCCATTTATTTAAAGCAAAAGTTATCGATCTTGATCTAAAAAACTT 1150
 384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
 1151 TGGGCCGGAACAGACGCTGGAGAACTTTGTGAAAGGCTCTTATGCTTATG 1200
 401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluGly 417
 1201 AAAGTTATGTAGAAATCCAGAACCAAGAGAAATCATAGATGAAGA 1250
 417 uGlyThrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
 1251 AGGTTGTTGCACACAGAGATATTGGGTATTACGATGAAGAAAAACATT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
 1301 TCTTTATCGTGAGATCGTTTGAAGCTTTAATCAATACAAAGATATACAA 1350
 451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePhe 467
 1351 GTACACACTGCTGTAATTAAGATCTGTTCTTTGCAACATCCAAATATTTT 1400
 467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLeuLeuProG 484
 1401 TGATCGCGCGCTGCTGCTGCTCCAGATCCTATAGCTGGGAGCTTCGG 1450
 484 lValAlaValValLeuLeuLysGlyLysSerMetThrGlyLysGluVal 500
 1451 GAGCTGTTGTTGTTACTTAAGAAAGAAATCTATGATGAAAAAAGAGTA 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyGly 517
 1501 ATGATATTACGTTGCTTACTCAAGTTTCAAATGCAAAACGTTTGGCTGTGG 1550
 517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
 1551 TGTCCGTTTGTGACGAGTACTAAAGGTCTCACAGTGAATAAATTGACG 1600
 534 lYysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
 1601 GTAAAGCAATTTAGAAATTAAGTGAAGAAACCAAGTGTGTAAGATG 1644
 seq_name: /SIDsz/gcdata/geneseq/geneseqn/NAI999.DAT:AAx84395
 seq_documentation_block:
 ID: AAx84395 standard; DNA; 1644 bp.
 XX
 AC: AAx84395;
 XX
 DT: 09-SEP-1999 (first entry)
 XX
 DE: L. lateralis luciferase coding sequence.
 XX
 KW: Luciferase; surfactant-tolerant; firefly; intracellular ATP assay; ss.
 XX
 OS: Luciola lateralis.
 XX
 PN: W09933997-A1.
 XX
 XX: 08-JUL-1999.
 PD: 24-DEC-1998; 98WO-JP05864.
 PF: XX

PR 26-DEC-1997; 97JP-0361022.

XX (KIKK) KIKKOMAN CORP.

XX Hattori N, Murakami S;

XX WPI; 1999-419109/35.

DR P-PSDB; AAY22182.

XX Modified luciferase tolerant to surfactants and useful for assay of

PT intracellular ATP

XX Example 2; Page 34-39; 56pp; English.

XX This sequence encodes a luciola lateralis luciferase of the invention.
 CC The protein is a surfactant-tolerant luciferase which is derived from
 CC firefly luciferase by replacement of the glutamic acid residue at
 CC position 490 by another amino acid. The luciferase can be used in an
 CC assay of intracellular ATP in biological samples containing cells, for
 CC example for establishing the presence of cells, and their quantity, in
 CC the sample. The method can be applied to foodstuffs, drugs,
 CC agrochemicals, beverages and environmental samples. The activity of the
 CC surfactant-tolerant luciferase is not lowered by the presence of high
 CC concentrations of surfactant and therefore an assay of high accuracy can
 CC be carried out in a simple manner.

XX Sequence 1644 BP; 528 A; 261 C; 349 G; 506 T; 0 other;

alignment_scores:

Quality: 2821.00 Length: 548
 Ratio: 5.148 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.818

alignment_block:

US-09-581-241-6 x.AAX84395 ..

Align seg 1/1 to: AAX84395 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
 1 ATGGAAACATGAGAGACGATGAAATATGTTGATGCTGCTGAACCATTT 50
 17 eTyrProIleGluGlySerAlaGlyAlaGluLeuArgLysTyrMetA 34
 51 TTACCTATTTGAAGAGGATGCTGAGACACAAATTGGCAAGATATAGG 100
 34 sPArGtYrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
 101 ATCGATATCAAAACTGAGCAATTGCTTTACTACGCACTTACCGGT 150
 51 ValAspTyrThrTyrAlaGluTyrLeuGlyLysSerCysCysLeuGlyG 67
 151 GTCGATTATACGATACCGCAATCTAGAAAAATCATGCTGCTCTAGACA 200
 67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
 201 GGCTTTAAAGATTATGTTGGTTTGTATGATGAGCAATTCGCTTATGCA 250
 84 eRgLuAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
 251 GTGAAAACGTGAGAAATCTTATTCCTGATTTACCGGTTTATTATA 300
 101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
 301 GGTGCTGGTGGTCCAACTATGAGATTACACTCTACGTGATGATTTGGT 350
 117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
 351 TCACAGTTTGGCATCTCTAAGCCAACTATTGATTTAGTTCTAAAAAAG 400
 134 LysLeuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150

401 GATTAGATAAAGTATTAACTGTACAAAAAAGCGTAAGTCTATTATAAAC 450
 151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
 451 ATGTATTATGTGACAGCAAGTGGATTATAGAGTTATCAATTCATGCA 500
 167 PAsnPheIleLysAsnThrProGlnGlyPheLysGlySerPheIle 184
 501 CACTTTATTAAAAAACCTCCACAAAGTTTCAAGATCAAGTATTA 550
 184 ySThrValGluValAsnArgLysGluValAlaLeuIleMetLysSer 200
 551 AAACGTAGAGTTAACCCGCAAAAGAAAGTGTCTCTTATATGAACCTC 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnI 217
 601 TCGGGTTCAACCGGTTGGCAAAAGGTGTGCACTTCACTACATAAATT 650
 217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
 651 GTCACACTAGATTTCACAGCTAGACATCCAAATTATGGAACCAAGTTT 700
 234 eRProGlyThrAlaIleLeuThrValValProPheHisHsGlyPheGly 250
 701 CACCAGCAGCGCTATTATTAACTGTAGTACCATTCATCATGATGTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
 751 ATGTTTACTACTTATGAGCTTACTTACTTGTGTTTCGTTATGTCATGTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGluAsnProLysC 284
 801 AACGAATTTGACAGAAAGACCTTTTAAAAACATGCAAGATTACAAAT 850
 284 ySserSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 851 GTTCACGCGTATTCTTGTACCGACCTTGTTCATTTCAATTTCAATTAAG 900
 301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleLeuAsn 317
 901 GAATTACTCGATAAATATGATTTATCAAAATTGTTAGTGAATTCATCTGG 950
 317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
 951 CGGACACCTTTATCTAAAGAAATGGTGAAAGCTGTGCTAGACGTTT 1000
 334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAla 350
 1001 ATTTACCGGCGTGTCTCAAGGCTATGTTTACAGAAACAACTCTGCA 1050
 351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
 1051 ATTATATACACCGGAAGGAGATGATTAACACGAGCTCTCTGCAAAAGT 1100
 367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThr 384
 1101 TGTGCCATTATTTAAAGCAAAAGTATCGATCTGATCAATAAAAAAAGCT 1150
 384 euGlyProAsnArgArgGlyGlyValCysValLysGlyProMetLeuMet 400
 1151 TGGGCCCAACAGCAGCTGAGAGAGTTTGTGTAAAGGCTCATATGCTTATG 1200
 401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
 1201 AAAGTTATGTAGATTAATCCAGAACCAAGCAAGAAATCATATGATGAAGA 1250
 417 uGlyThrPheLysHisThrGlyAspIleGlyTyrThrAspGluGlyHisP 434
 1251 AGCTTGTGTCACACAGAGATATTGGGTATTAGCATGAAGAAAAACATT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
 1301 TCTTTATCGTGAATGCTTTGAAGTCTTTATCAATTAACAAAGGATATCAA 1350


```

|||||
751 ATGTTTACTACTTAGGCTATCTACTTGCTTGGTATGTCAGTCTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
|||||
801 AACGAAATTTGACGAAGAGACTTTTAAAAACACTGCACAAATACAAAT 850
284 ySseSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
|||||
851 GTTCAGAGGTTATCTGTGTACCGACTTGTGTGCAATCTTAATAGAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
|||||
901 GAATTACTCGAATTAATATGATTTACAAATTTAGTGAATTCATCGACG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
|||||
951 CGAGACACCTTATCTAAAGAAATTTGGTGAAGCTGTGCTAGAGCTTTTA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
|||||
1001 ATTATCCGGGTGTCTCAAGGCTATGTTTACAGAAACAACTCTGCA 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
|||||
1051 ATTATATACACACCGAAGGCGATGATTAACAGAGTCTCTGGCAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
|||||
1101 TGTGCATTTATTTAAAGCAAAAGTTATGATCTGATCTTAAAAAACTT 1150
384 euGlyProAsnArgArgGlyGlyValLysValLysGlyProMetLeuMet 400
|||||
1151 TGGGCCCGACAGACGTGAGAAAGTTTGTGTAAGGCTCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluG 417
|||||
1201 AAAGGTATGTAGATATATCCAGAACCAACAGAAATCATGATGAGAAGA 1250
417 uGlyTyrPheLysThrGlyAspIleGlyTyrTyrAspGluGlyLysHisP 434
|||||
1251 AGCTTGGTGGCACACAGAGATATTTGGTATTCATGATGAAGAAACAT 1300
434 hepPheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
|||||
1301 TCTTTATCTGTGATGCTTTGAAGTCTTTATACAAATACAAAGATATCA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
|||||
1351 GTACACACTGCTGAATTAAGATCTGTTCTTTGCAACATCCAAATATTTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLuleuProG 484
|||||
1401 TGAATGCCGGCGTGTGCGCTGCCATCTTAAGCTGAGTTCGCGG 1450
484 lValAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
|||||
1451 GAGCTCTTGTGTACTTAAGAAAGAAATCTATGACTGAAAGAAAGCTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyL 517
|||||
1501 ATGATTTAGCTGTGCTAGTCAAGTTTCAAAATGCAAAAACGTTTGGTGTG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
|||||
1551 TGTCCGTTTGTGACGAAGTACTTAAAGTCTCACTGCTAAATTTGAGCG 1600
534 lYrsAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
|||||
1601 GTAAAGCAATTTAGAAATATCTGAAAGAAACCACTTCTTAAGATG 1644
seq_name: /SIDSZ/gcgdata/geneseq/geneseqn/NA1993.DAT:AA034745

```

```

seq_documentation_block:
ID   AA034745 standard; cDNA to mRNA; 1644 BP.
XX
XX
AC   AA034745;
XX
DT   24-MAY-1993 (first entry)
XX
DE   Thermostable luciferase (wild-type).
XX
KW   Firefly; Luciola cruciata; GENUI; Luciola lateralis; HEIKE;
KW   mutant; assay; ss.
XX
OS   Luciola lateralis.
XX
PN   EP524448-A.
XX
PD   27-JAN-1993.
XX
PF   26-JUN-1992; 92EP-0110808.
XX
PR   27-JUN-1991; 91JP-0157117.
PR   29-NOV-1991; 91JP-0317064.
PR   22-MAY-1992; 92JP-0131057.
XX
XX   (KIKK ) KIKKOMAN CORP.
XX
PI   Eliehl N, Naoki K;
XX
DR   WPI; 1993-028553/04.
DR   P-PDB; AAR30803.
XX
PT   New DNA sequence of a thermostable luciferase - has the aminoacid
PT   at position-217 of Luciola cruciata or lateralis luciferase
PT   replaced by a hydrophobic aminoacid
XX
PS   Claim 1-3; Page 29-30 + 21-24; 33pp; English.
XX
XX   A DNA sequence of a thermostable luciferase of a firefly, which
CC   encodes the amino acid sequence of a wild-type firefly luciferase in
CC   which an amino acid at the 217 position or an amino acid at the
CC   position equiv. to the 217 position of the luciferase of Luciola
CC   cruciata (GENUI firefly) of Luciola lateralis (HEIKE firefly) is
CC   replaced by a hydrophobic amino acids, esp. Ile, Leu or Val, is
CC   claimed. The mutant luciferase is identical in properties to the
CC   wild-type luciferase except that it is stable when heated to high
CC   temps., e.g. 50 degrees C. The luciferase can be used in assays for
CC   e.g. ATP.
XX
SQ   Sequence 1644 BP; 529 A; 262 C; 349 G; 504 T; 0 other:
XX

alignment_scores:
Quality: 2814.00      Length: 548
Ratio: 5.144          Gaps: 0
Percent Similarity: 99.818      Percent Identity: 99.635

alignment_block:
US-09-581-241-6 x AA034745 ..

Align seg 1/1 to: AA034745 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValIleValTyrGlyProGluProPh 17
|||||
1 ATGCAAAACATGACGAMCGATGAAATATTTGTATGCTGCTGCAACCATTT 50
|||||
17 eTyrProIleGluGluGlySerAlaGlyAlaGluLeuArgTyrTyrMetA 34
|||||
51 TTACCTTATTGAAGGCGATCTGCTGGACACAAATGCGCAAGTATATNG 100
|||||
34 sParGTYrAlaLysLeuGluAlaIleAlaPheThrAsnAlaLeuThrGly 50
|||||
101 ATCGATATGCAAAACTTGACGCAATTCCTTTACTAAGCACTTACCGGT 150

```

[illegible]


```

401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleLeuLeuGlu 417
    |||||
1201 AAGGTTATATGATATATCCAGAACGACAGAGAAATCATAGATGAGAGA 1250
    |||||
417 uGlyTyrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluHisP 434
    |||||
1251 AGGTTGGTTCACACAGAGATATGGGTATTCAGATGAGAAACATTT 1300
    |||||
434 hePheIleValAspArgIleuLysSerLeuIleuLysTyrGln 450
    |||||
1301 TCTTTATCGGATCGTTGAACTTTAATCAATAACAAAGATATCAA 1350
    |||||
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIleP 467
    |||||
1351 GTACCCACGTGCTGATTAATCTGTTCTTTCACACATCCAAATATTTT 1400
    |||||
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
    |||||
1401 TGATGCCGGCGTTCGGCGCTCCAGATCTATAGCTGTGAGCTTCCGG 1450
    |||||
484 IValAlaValAlaValLeuLysGlyLysSerMetThrGluLysGluVal 500
    |||||
1451 GAGCTGTGTGACTTGAATAAAGAAATCTATGACTGAAAGAAAGTAA 1500
    |||||
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
    |||||
1501 AAGGATTAAGTTCGTAGTCAGTTCAATGCAAAAGCTTGGGTGCTGG 1550
    |||||
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
    |||||
1551 TCTCGCTTTTGGAGGAAAGTACCTTAAGAGTCTCACTGTAAATTTGAGC 1600
    |||||
534 IValAlaIleArgGluIleLeuLysLysProValAlaLys 547
    |||||
1601 GTAAGCAATTAAGAAATACGAAAGAAACAGTTGCTAAG 1641
    |||||
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NAL198.DAT:AAV32467
seq_documentation_block:
ID AAV32467 standard; DNA: 1920 BP.
XX
AC AAV32467;
XX
DT 10-SEP-1998 (first entry)
XX
DE Luciferase-lysoastaphin fused gene.
XX
KW Luciferase-lysoastaphin fused gene; chimeric; bioluminescence; ds.
XX
OS Chimeric - Staphylococcus simulans.
XX
OS Chimeric - Luciola lateralis.
XX
FH Key
FT Location/Qualifiers
FT CDS 1..1920
FT /*tag= a
FT /product= "Luciferase-lysoastaphin fusion protein"
FT /note= "Nucleotides 1-1647 encode luciferase
FT protein while nucleotides 1648-1920 encode
FT lysoastaphin; CDS does not contain stop codon"
XX
PN JP10150991-A.
XX
PD 09-JUN-1998.
XX
PE 25-NOV-1996; 96JP-0328042.
XX
PR 25-NOV-1996; 96JP-0328042.
XX
PA (KIKK ) KIKKOMAN CORP.
XX
DR WPI; 1998-379994/33.
DR P-PSDB; AAW48856.

```

```

XX
PT New Luciferase-lysoastaphin fused protein - useful for
PT bioluminescence analysis
XX
PS Disclosure; Pages 6-7; 10pp: Japanese.
XX
CC The invention claims for a luciferase-lysoastaphin fused gene which
CC encodes a fused protein in which a peptide part consisting of
CC luciferase, from Luciola lateralis, is connected to a peptide part
CC consisting of lysoastaphin from Staphylococcus simulans. The method
CC described by the invention can be used to prepare a luciferase-
CC lysoastaphin fused protein efficiently. The fused protein is useful
CC for bioluminescence analysis.
XX
SQ Sequence 1920 BP; 627 A; 310 C; 407 G; 576 T; 0 other;

alignment_scores:
    quality: 2812.00      length: 547
    ratio: 5.141         gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.634

alignment_block:
US-09-581-241-6 x AAV32467 ..

Align seg 1/1 to: AAV32467 from: 1 to: 1920

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
    |||||
1 ATGGAAGAACATGGAGAACATGAAATATTTGTTATGATGCTGAACCAAT 50
    |||||
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetA 34
    |||||
51 TTACCTTATTGAAGAGGATCTGCTGACACAAATTCGCAAGTATATGG 100
    |||||
34 sPArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
    |||||
101 ATCGATATGCAAAACTTGAGCAATTCGTTTACTTAACGACACTTACCG 150
    |||||
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysGlyLeuGly 67
    |||||
151 GTGCAATTATACGACCGCAATCTTAGAAATCATGCTGTCTAGAGAA 200
    |||||
67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuCys 84
    |||||
201 GCGTTTAAAGAAATTAAGTTGCTGTTGATGAGAAATTCGCTATATGA 250
    |||||
84 eGluAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
    |||||
251 GTGAAACTGTGAAAGATCTTATTCGTATTAGCCGTTTATTTATA 300
    |||||
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
    |||||
301 GGTGTCGGTGGTCCCAACTAATAGATTTACCTTACGTAAGATTGGT 350
    |||||
117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
    |||||
351 TCACAGTTAGGCATCTCTAAGCCAACATTTGATTTAGTTTAAAGAA 400
    |||||
134 ILeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
    |||||
401 GATTAGATTAAGTTATTAAGTACAAAAAGCGTAACGTATTAAAC 450
    |||||
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
    |||||
451 ATTGTTATATTGGACAGCAAGTGAATATAGAGTTATCAATTCATGGA 500
    |||||
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
    |||||
501 CAACCTTTATTAAAAAAACACTCCACAAGGTTTCAAGATCAACGTTTAA 550
    |||||
184 ySThrValGluValAsnArgLysGluGlnValAlaLeuIleMetLysSer 200
    |||||

```

551 AAACGTAGAGTTAAACCGCAAGAACAAAGTGCCTATATGACCTC 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnL 217
 601 TCGGGTTCACACCGGTTGGCAAAAGGTGCAACTTACTCATGAAATTT 650
 217 eValThrArgPheSerHisAlaArgAspProLeyGlyAsnGluValS 234
 651 GGTCACCGGTTTTCTCACGCTAGAGATCCCAATTTATGAAACCAAGTT 700
 234 eProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
 701 CACACGACGCGGTATTTAACTGTAAGTACATCCATCATGTTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMet 267
 751 ATGTTTACTACTTTAGCTATCTAACTGGTGGTTTGTATGTCAGATT 800
 267 uThrLysPheAspGluGlnThrPheLeuLysThrLeuGlnAspTyrLysC 284
 801 AACGAAATTTGACGACAGACACTTTTAAAAACACTGCACAGATTACAA 850
 284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 851 GTTCACGCGTATTTCTGACCGCACTTGTTCGAATTTCTTAATAGAGT 900
 301 GluLeuAsnAspLysTyrAspLeuSerAsnLeuValGlnIleAlaSerG 317
 901 GAATTTACTCGTAAATATGATTTTCAAAATTTAGTTAAATTTGCATCTG 950
 317 yGlyAlaProLeuSerLysGlnIleGlyGluAlaValAlaArgArgPhe 334
 951 CGGAGCACTTATCTAAAGAAATTTGGTGAAGCTTGCTAGACGTTTAA 1000
 334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAla 350
 1001 ATTTACCGGCGTTCGTCACGCTATGTTTAAACGAAACCTCTGCA 1050
 351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
 1051 ATTATTTATCACACCGAGGCGGATGATTAACCAAGTCTTGTGCAAAAGT 1100
 367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
 1101 TGTGCATTATTTAAAGCAAAAGTTATGATCTGATCTAAAAAAAAGTT 1150
 384 euGlyProAsnArgArgGlyGlyValCysValLysGlyProMetLeuMet 400
 1151 TGGGCCCGACACAGACGTGAGAGAGTTTGTGTAAGGCGTCTATGCTTATG 1200
 401 LysGlyTyrValAlaAspAsnProGluAlaThrArgGluIleIleAspGluG 417
 1201 AAAGTTTATGTGATATATCCAGAACCAACAGAAATCATAGATGAAGA 1250
 417 uGlyTyrPheHisThrGlyAspIleGlyTyrTyrAspGluLysHisP 434
 1251 AGCTGTGTTGCACACAGAGATATGGGTATATGATGATCAAAAAACAATT 1300
 434 hepPheIleAlaAspArgLeuLysSerLeuIleLysTyrGlyGlyTyrGln 450
 1301 TCTTTATGCTGATGCTTGAAGTCTTTAATCAAAATACAAAGGATATACA 1350
 451 ValProProAlaGluLeuGlnLysSerValLeuLeuGlnHisProAsnIlePh 467
 1351 GTRACACCGCTGATTAAGATCTGTTCTTTGCAACAATCCAAATATTTT 1400
 467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
 1401 TGAATGCCGCGCTTGGCGGTTCAGATCCATTAAGTGTGAGCTTCCGG 1450
 484 lValAlaValAlaValLeuLysGlyLysSerMetThrGluLysGluVal 500
 1451 GAGCTTTGTTGTTACTGAAAAAGGAAATCTATGACTGAAAAAGACAGTA 1500

501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
 1501 ATGATATTACGTTGCTAGTCACTTTCAATGCAAAACGTTCCGGTGG 1550
 517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
 1551 TGCCTGTTTGTGACGAGTACTTAAAGTCTCACTGCTAATAATTTGACG 1600
 534 lYlValAlaIleArgGlnIleLeuLysLysProValAlaLys 547
 1601 GTAAGCAATTTAGAGAAATATCTGAAGAAACCAAGTTGCTAAG 1641
 seq_name: /SID52/gcgdata/geneseq/NA1990.DAT:AA003257
 seq_documentation_block:
 ID AA003257 standard; DNA; 1644 BP.

AA003257:
 22-JUL-1990 (first entry)
 DE Luciola lateralis recombinant luciferase gene.
 KW Luciferase; enzyme; ATP; plasmid pHLf7; firefly; ss.
 OS Luciola lateralis.
 FH Key Location/Qualifiers
 FT CDS 1..1644
 FT /*tag= a
 FT /product= luciferase
 FT /EC_number= EC-1.13.12.7

EP353464-A.
 PD 07-FEB-1990.
 PF 30-JUN-1989; 89EP-0111958.
 PR 22-DEC-1988; 88JP-0162402.
 PR 01-JUL-1988; 88JP-0162402.
 XX PA (KIRK) KIRKOMAN CORP.
 XX PI
 XX Tatsumi H, KajiYama N, Nakano E;
 DR WPI: 1990-038240/06.
 DR P-PSDB; AAR03731.
 XX
 PT New gene encoding luciferase
 PT derived from Luciola lateralis and expressed in and purified
 PT from Escherichia coli genus cells.
 XX
 PS Claim 3; page 24; 43pp; English.
 CC This enzyme is useful for determining ATP levels. It is
 CC contained in plasmid pHLf7.
 XX
 SO Sequence 1644 BP; 529 A; 264 C; 348 G; 503 T; 0 other;

alignment_scores:
 quality: 2811.00 length: 548
 Ratio: 5.139 Gaps: 0
 Percent Similarity: 99.818 Percent Identity: 99.453

alignment_block:

US-09-581-241-6 x AA003257 ..

Align seg 1/1 to: AA003257 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17

```

|||||
1  ATGGAAAAATGGAGAACGATGAAATATGTGTATGCTCGAACCAT 50
17  eHyrrProIIeGluGluGlySerIaGlyAaGlnLeuAArgLysTyrMet 34
51  TTAACCTATGTAAGAGAGATCTGCTGAGCACAATTGCCAAGTAATATG 100
34  sPaRgtYrAlaYsLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101  ATCGATACGCCAAACTTGGAGCAATTCCTTTACTACCGACTTACC 150
51  ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
151  GTCGATTATACGACGCCGCAATCTTAGAAAAATCATGCTGCTAGAGA 200
67  uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
201  GGCTTAAAGAAATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 250
84  eRgLUAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPhe 100
251  GTGAAAACTGTGAAGAAATCTTATCTCTGTATACCGCGTTATTATTA 300
101  GYValGlyValAlaProThrAsnGluIleTyrThrLeuAArgLysLeu 117
301  GGTGTCGGGTGCTCCACACTATGATTAACACTACGTAAGTAATGCT 350
117  lHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
351  TCACAGTTAGCATCTCTAAGCAACAATGTATTTAGTTCTTAAAAAG 400
134  lYLeuAspLysValIleThrValGlnLysThrValIleAlaIleTyr 150
401  GATTGATTAAGTATTAACCTGTACAAAAAAGCGTAACGCTATTAAAC 450
151  lIleValIleLeuAspSerLysValAspTyrArgIleTyrGlnSerMet 167
451  ATTGTTATATGAGACGCAAACTGATTTATGAGTTATCAATCCATGCA 500
167  PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerP 184
501  CAACCTTATTAATAAAAAACACTCCACAAGGTTTCAAGGATCAAGTT 550
184  YsThrValGlnValAsnArgLysGluGlnValAlaIleuIleMetS 200
551  AAACGTGTAAGATTAAACCCAAAGACAAAGTGTCTTATTAATGAAC 600
201  SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAs 217
601  TCGGGTCAACCGGTTTGCCTAAAGTGTGCAACTTACTCATGAATA 650
217  eValThrArgPheSerHisAlaArgAspProIleTyrGlnGlnValS 234
651  AGTCACCTAGATTCTTCACGCTAGAGATCAATTTATGAAACCAAGTT 700
234  ePProGlyThrAlaIleLeuThrValValIleProPheHisGlyPhe 250
701  CACACGAGCAGCGTATTTTAAGTATGATCAATCCATCATGCTTTG 750
251  MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMet 267
751  ATGTTTACTACTTATGAGCTATCTAATCTGTGTTTCTGTAATGCA 800
267  uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyr 284
801  AACGAAATTTGACGAAAGAGACTTTTAAACACATGCAAGATTACA 850
284  YsSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArg 300
851  GTTCAACGCTATATCTGTACGACTTGTTCGATTCCTTAATGAAGT 900
301  GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSer 317
|||||

```

```

901  GAATTAATCGATAAATATGATTTATCAAAATTTAGTTGAATTCACATCTG 950
317  YGlyAlaProLeuSerLysGluIleGlyAlaValAlaArgArgPhe 334
951  CGGAGCACCTTTATCTAAAGAAATGCTGAAGCTGTTGCTAGACGTTT 1000
334  snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
1001  ATTTACCGGGGTGTTGCTCAAGGCTATGCTTTAACACAAACACCTCTG 1050
351  lIleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLys 367
1051  ATTATTTATACACACCGGAGCGATGATFAAACCGAGTCTTGCAAAAG 1100
367  lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThr 384
1101  TGTCCATTATTTAAAGCAAAAGTATGATCTTGTACTTAAAAAACTT 1150
384  euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
1151  TGGGCCCGAACAGACCTGGAGAAAGTTTGTGTAAAGGCTCTTATGTT 1200
401  LysGlyTyrValAspAsnProGluAlaThrArgGluIleLeuAspGlu 417
1201  AAAGTTATGTAGATTAATCCAGAACCAACAGAAATCATAGATGATAGA 1250
417  uGlyThrPheHisThrGlyAspIleGlyTyrTyrAspGluGluLysHis 434
1251  AGTTTGTTGTCACACAGAGATATTTGGTATTAACGATGAAGAAAACT 1300
434  hPheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301  TCTTATCTGATGCTTGAAGTCTTTAAATCAAAATACAAAGATATCA 1350
451  ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIle 467
1351  GTACCACTGCTGAATTAAGATCTGTTCTTTGCAACATCCAAATATTT 1400
467  eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuPro 484
1401  TGATGCCGCGCTGCTGGGCTCCAGATCCTTACTGCTGAGCTTCCGG 1450
484  lYAlaValValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
1451  GAGCTGTTCTGTTACTTGAAAAAGGAAATCTATGACGAAAAAGAGTA 1500
501  MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
1501  ATGGATTACGTTGCTAGTCACAGTTTCAATGCAAAACGTTGCGGTGG 1550
517  YValArgPheValAspGluValProLysGlyLeuThrGlyLysLeuAsp 534
1551  TGTCCGTTTGTGGAGCAATCTAAGAGTCTCACTGCTGTAATGTAAG 1600
534  lYsAlaIleAlaArgGluIleLeuLysLysProValAlaLysMet 548
1601  GTAAAGCAATTAGAGAAATTAAGTGAAGAAACAGTGTCAAGATG 1644
seq_name: /SID52/gcgdata/geneseq/geneseqn/NA1997.DAT:AA63268
seq_documentation block:
ID  AA63268 standard; DNA; 1704 BP.
XX
AC  AA63268;
XX
DT  20-MAY-1997 (first entry)
XX
DE  Firefly luciferase conjugated with 23 aa peptide at 5' end.
XX
KW  Fusion protein; firefly; luciola lateralis; luciferase; biotinylation;
KW  wild type; E.coli; bioluminescence assay; quantification; ligand;
KW  receptor; ds.
XX

```



```

1167 ATTATTAAAGCAAAAGTTATCGATCTTGATCTTAAACCAAACTTTGGGCC 1216
386 rosmnprgargylgluvalcysvallysglyprometleumetlysgly 402
1217 CGAACGACGCTGGAGAGATTGTGTAAAGGGCTCTATGCTTATGAAGGT 1266
403 TyrValAspAsnProGluValThrArgGluIleLeuAspGluGlyTr 419
1267 TATGTAGATTAATCCAGAACACAGCAAGAAATCATAGATCAAGAGCTTG 1316
419 pleuHstHnglyAspIleGlyTyrTrpAspGluGlyLysHisPhePhe 436
1317 GTTGCACACAGGAGATTTGGTATTACGATGAAGAAACAAATTTCTTTA 1366
436 leValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGluValPro 452
1367 TCGTGATCGTTGAAGTCTTTATCATCAATACAAAGCATATCAAGTACCA 1416
453 ProAlaGluLeuGluSerValLeuLeuGluHisProAsnIlePheAspAl 469
1417 CCGCGAATTAAGAAATCTGCTTTTGCACCAATCAATATTTTGATGC 1466
469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyValay 486
1467 CCGCGTGTGGCGTTCAGATCTTATAGCTGGTGAAGCTTCCGGAGCTG 1516
486 aValValLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
1517 TTGTTGTAATCTGAAAAAGAAATCATGACAGTGAAGAGTAATGCAAT 1566
503 TyrValAlaSerGluValSerAsnAlaLysArgLeuArgGlyGlyValar 519
1567 TACGTTGCTAGTCAAGTTCAATGCAAAACGTTGCGTGGTGGTCCG 1616
519 gpHeValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536
1617 TTTTGTGACGAAGTACCTAAAGGTCTGCTGTAATTTGACGCGTAAG 1666
536 laIleArgGluIleLeuLysLysProValAlaLysMet 548
1667 CAATTAGAGAAATACTGAAGAAACGTTGCTTAAGATG 1704

seq_name: /SIDS2/gcgdata/geneseq/NA1998.DAT:AAV23595
seq_documentation_block:
ID AAV23595 standard; cDNA to mRNA; 1704 BP.
XX
AC AAV23595;
XX
DT 16-JUL-1998 (first entry)
XX
DE Antibody-Firefly luciferase fusion protein gene.
XX
KM Firefly luciferase; antibody-luciferase fusion protein; ds.
XX
OS Luciola cruciata.
XX
FH Key location/Qualifiers
FT CDS 1..1704
FT FT /*tag= a
FT FT /note= "no stop codon given"
XX
XX JP09187281-A.
XX
XX 22-JUL-1997.
XX
XX 09-JAN-1996; 96JP-0001812.
XX
XX 09-JAN-1996; 96JP-0001812.
XX
XX (KIKK ) KIKKOMAN CORP.
XX
XX WPI; 1998-275089/25.

```

```

DR P-PSDB; AAV53883.
XX
XX Antibody-firefly luciferase fused protein - and related products
PT i.e. firefly luciferase fused gene, recombinant DNA and its
PT preparation
XX
XX PS Disclosure: Page 10-11; 17pp; Japanese.
XX
XX This sequence encodes a fusion protein of the invention. The protein is
CC a antibody-firefly luciferase fusion protein, in which an antibody part
CC consisting of a peptide having antibody activity is combined with an
CC enzyme part consisting of firefly luciferase.
XX
XX SQ Sequence 1704 BP; 541 A; 275 C; 361 G; 527 T; 0 other:

alignment_scores:
Quality: 2799.00 Length: 546
Ratio: 5.126 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.267

alignment_block:
US-09-581-241-6 x AAV23595 ..
Align seg 1/1 to: AAV23595 from: 1 to: 1704

3 AsnMetGluAsnAspGluAsnIleValIleTyrGlyProGluProPheTyrPr 19
.....
67 AGCTCGAGACAGATGAAAAATATGTGTATGCTCGAAGCAATTTTACC 116
19 orleGluGluGlySerAlaGlyValAlaGluLeuArgLysTyrMetAspArgT 36
117 TATTAAGAGGGATCGCTGGAGACAAATTCGCCAAGTATATGATGATGAT 166
36 yAlAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValaAsp 52
167 ATGCAAACTTGAGCAATTCCTTTACTTAACGCACTTACCGGTGTCGAT 216
53 TyrThrTyrAlaGluTyrLeuGluLysSerCysGlyLeuGluAlaIle 69
217 TATACGACCGCGAATACTTACAAAAATATCTGCTTACGAGAGGCTTT 266
69 uLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuCysSerGluA 86
267 AAAGAAATTATGTTGTTGTTGATGAGAAAGATTCGTTATATCAGAGAA 316
86 snCysGluGluPhePheIleProValIleAlaGlyLeuPheIleGlyVal 102
317 ACTGTGAAGAGTTCTTTATCTCTGATTAAGCCGGTTTATTTAAGGTGTC 366
103 GlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHisSe 119
367 GGTGTGGCTCCCAACTAATGAGATTACACTGACTGATTTGTTACACAG 416
119 rLeuGlyIleSerLysProThrIleValIlePheSerSerLysGlyLeuA 136
417 TTTAAGCATCTCTTAAGCCAAATGTATTTAGTTCTAAAAAAGGATTAG 466
136 sPlyValAlleThrValGluLysThrValIleAlaIleLysThrIleVal 152
467 ATTAAGTATTAATCTACAAAAAACGTAAGTACTGCTATTTAAACCATGTT 516
153 lleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAspAsnph 169
517 ATATTGACAGCAAAAGTGAATTAAGAGTTATCAATCCATGAGACAACTT 566
169 eileLysLysAsnThrProGluGlyPheLysGlySerSerPheLysThrV 186
567 TATTAAAAAAACACTCCACAAGGTTTCAAAAGATCAAGTTTAAACTG 616
186 aIleGluValAsnArgLysGluGluValAlaLeuIleMetAsnSerSerGly 202
617 TAGAAGTTAACCGCAAGACAGTTCCTTATATATGAACTCTTCGGGT 666

```

203 SerThrGlyLeuProIysGlyValGlnLeuThrHisGluAsnIleValIle 219
 |||||
 667 TCACCCGGTTGGCCAAAAGGTGTGCACTTACTCATGAAAATTTGGTCAC 716
 219 rArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValSerPro 236
 |||||
 717 GCGTTTTCACGCTAGAGATCCAAATTTATGGAAACCAAGTTTCACGAC 766
 236 lYThrAlaIleLeuThrValValProPheHisHisGlyPheGlyMetPhe 252
 |||||
 767 GCAGCGCTATTTTACTGTAGTACCATTCATCATGTTTGGATGTTT 816
 253 ThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLeuThr 269
 |||||
 817 ACTACTTATGAGCTATCTAACTGTGTGTTCTGATCTCATGTTAACGAA 866
 269 sPheAspGlnGluThrPheLeuLysThrLeuGlnAspTyrLysCysSer 286
 |||||
 867 ATTTGACGAGAGACTTTTAAAACACCTGCAAGATTACCAATGTTCAA 916
 286 eValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
 |||||
 917 GCGTTATCTTGTACCGACTTGTGTTGCAATTTCTTAATAGAAAGTAAT 966
 303 LeuAspLysTyrAspLeuSerAsnLeuValGlnIleAsnGlyGlyAla 319
 |||||
 967 CTGCTAAATATGATTTATCAAAATTTAGTGAATTCATCTGCGGAGCC 1016
 319 aProLeuSerLysGluIleGlyGluAlaValAlaArgAspPheAsnLeu 336
 |||||
 1017 ACCTTATCTTAAAGAAATTTGTTGAAGCTGTGTGACGTTTATTTAT 1066
 336 rGlyValAlaGlnGlyTyrGlyLeuThrGluThrSerAlaIleIle 352
 |||||
 1067 CCGGTGTCGTCACAGGCTATGTTTAAACAGAAACCTCTGCAATTAAT 1116
 353 lIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysValValPr 369
 |||||
 1117 ATCAACCCGGAAGGCGATTAACCAAGTGTCTGTGCAAAAGTTGTGCC 1166
 369 oLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP 386
 |||||
 1167 ATTATTTTAAAGCAAAAGTTATCGATCTTGATCTAAATAAACTTTGGGCC 1216
 386 rAsnArgArgGlyGlyValLysValLysGlyProMetLeuMetLysGly 402
 |||||
 1217 CCAACAGACGTGAGAGAGTTGTGTAAAGGCTCTATCTTATGAAAGCT 1266
 403 TyrValAspAsnProGluAlaThrArgIleIleAspGlnGluGlyTr 419
 |||||
 1267 TATGTAGATTAATCCAGAAACCAAGAAATCATAGATGAAGAAAGGTTG 1316
 419 pLeuHisThrGlyAspIleGlyTyrTyrAspGlnGluLysHisPhePhe 436
 |||||
 1317 GTTGCACACAGAGATATTTGGTATTACGATGAAGAAAAACTTTCTTTA 1366
 436 lValAlaAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
 |||||
 1367 TCGTGATGCTTTGAAGCTTTTAAATCAATACAAAGGATATCAAGTACCA 1416
 453 ProAlaGlnLeuGlnSerValLeuLeuGlnHisProAsnIlePheAspAl 469
 |||||
 1417 CCTGCTGATTTAGAAATCTGTTTTCGCAACATCCAAATATTTTGATGCC 1466
 469 aGlyValAlaGlyValProAspProIleAlaGlyLeuLeuProGlyAlaVal 486
 |||||
 1467 CCGCGTTGCTGGGCTTCAGATCTATACCTGTGAGCTTCGCGGAGCTG 1516
 486 aValValLeuLeuLysGlyLysSerMetThrGluLysGluValMetAsp 502
 |||||
 1517 TTGTTGTACTGAAAAAGAAAAATCTATGACTGAAAAAGAAAGTAATGCA 1566

503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValAla 519
 |||||
 1567 TACGTTGCTAGTCACTTCAATGCAAAACGTTTCCGGTGGTGTCCG 1616
 519 pPheValAspGlnValProLysGlyLeuThrGlyLysIleAspGlyLysAla 536
 |||||
 1617 TTTTGTGACGAGATACCTAAAGGCTCTACTGCTAAATTTGACGGTAAAG 1666
 536 lAlaArgGlnLeuLeuLysLysProValAlaLysMet 548
 |||||
 1667 CAATTAGAGAAATACGAAAGAAACGTTGCTAAGATG 1704
 seq_name: /SID52/gcgdata/geneseq/geneseq/NA1996.DAT: AAT33850

seq_documentation_block:
 ID AAT33850 standard; DNA: 2019 BP.
 AC AAT33850;
 DT 12-NOV-1996 (first entry)
 DE DNA encoding streptavidin/luciferase fusion protein.
 KW Streptavidin; luciferase; fusion protein;
 KW Streptomyces avidinii; Luciola lateralis; firefly;
 KW recombinant production; industry; ds.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2019
 FT /*tag= a
 PN JP07289264-A.
 PD 07-NOV-1995.
 XX
 PF 27-APR-1994: 94JP-0090275.
 XX
 PR 27-APR-1994: 94JP-0090275.
 XX
 PA (KIRK) KIRKMAN CORP.
 XX
 DR WPI: 1996-015269/02.
 DR P-PSDB: AAW04208.
 XX
 PT New mutant streptavidin (SA) gene - fused to a firefly luciferase
 gene, for the recombinant prepn. of a SA-FL fused protein
 XX
 PS Example 1; Pages 5-6; 12pp; Japanese.
 XX
 CC The present sequence encodes a streptavidin/luciferase fusion
 protein, comprising the Streptomyces avidinii streptavidin gene
 and the Luciola lateralis (firefly) luciferase gene. The fusion
 protein can be prepd. by inserting the recombinant DNA encoding
 it into a Escherichia species microbe, culturing the transformed
 CC microbe in a medium and collecting the fusion protein from the
 culture. The fusion protein can be used in industry.
 CC
 XX
 SQ Sequence 2019 BP; 590 A; 412 C; 467 G; 550 T; 0 other:

alignment_scores:
 Quality: 2799.00 length: 546
 Ratio: 5.126 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.267

alignment_block:

US-09-581-241-6 x AAT33850 ..

Align seg 1/1 to: AAT33850 from: 1 to: 2019

3 AsnMetGluAsnAspGluAsnIleValIleValIleGlyProGluProPheTyrPr 19

```

382  TCCCTGAGACGATGAAATATTGTGTATGTCCTGACCATTTTACC 431
19  oIIeGIuGIuSeraIaGIuAgluInleuAArglySTyMetAspArgT 36
432  TATTGAGAGGAGATCTGCTGAGACACATTTGCCAAGTATATGATGAT 481
36  yralAlaLysleuGIuAlaIleAlaPheThrAsnAlaLeuThrGIuValAsp 52
482  ATGCAGAACTTGAGCAATTTGCTTTACTACCACTTACCGGTTCGAT 531
53  TyrThrTyrAlaGIuTyrleuGIuLysSerCysCysleuGIuAlaIle 69
532  TATACGTACGCCGAAATCTTAGAAAAATCATGCTGTACGAGAGGCTTT 581
69  uLysAsnTyrGIuLeuValValAspGIuArgIleAlaLeuCysSerGIuA 86
582  AAGAGATTATGTTTGTGTTGTTGATGCAAGCAATTTGCTTTCAGTGA 631
86  sncysGIuGIuPhePheIleProValleuAlaGIuLeuPheIleGIuVal 102
632  ACTGTCAAGAGTTCTTATTCCTGTATAGCCGTTTATTATTAGTGTC 681
103  GIuValAlaProThrAsnGIuIleTyrThrleuAArgGIuLeuValHisSe 119
682  GGTGTGCTCCACATATAGATTTACCTCTACGTAAGTAATGGTTCACAG 731
119  rleuGIuIleSerLysProThrIleValPheSerSerLysGIuLeuA 136
732  TTTAGCATCTCTAAGCCAAATTTGATTTAGTTCTTAAAAAGATTAG 781
136  sPLyValIleThrValGIuLysThrValThAlaIleLysThrIleVal 152
782  ATTAAGTTATTAAGTCAAAAAACGTAACGCTATTAATAAACCATTTGT 831
153  IleLeuAspSerLysValAspTyrArgIleTyrGIuInSerMetAspAsnPh 169
832  AATTTGAGACAGCAAGTGTATTTAGAGTTATCAATTCATGACACACTT 881
169  eIleLysLysAsnThrProGIuGIuPheLysGIuSerSerPheLysThrV 186
882  TATTAATAAAAAACACTCCACAGGTTTCAAGATCAAGATTTTAAAACTG 931
186  aIGluValAsnArgLysGIuInleuAlaIleuIleMetAsnSerSerGIy 202
932  TAGAAGTTAACCGCAAAAGACAAAGTTGCTCTTATTAATGAACCTTCGGGT 981
203  SerThrGIuLeuProLysGIuValGIuLeuThrHisGIuAsnIleValTh 219
982  TCAACCGGTTTGCCTCAAAAGGTGTGCAACTTACTCATGAATAATTTGGTCC 1031
219  rArgPheSerHisAlaArgAspProIleTyrGIuLysnGIuValSerProG 236
1032  GCGTTTCTCCAGCTAGATCAATTTATGAAACCAAGTTTTCACAG 1081
236  IYThrAlaIleLeuThrValValProPheHisHisGIuPheGIuMetPhe 252
1082  GCACGCGATTTTAACTAGTACCATTCATCATGTTTGGTATGTTT 1131
253  ThrThrleuGIuTyrleuThrCysGIuPheArgIleValMetleuThrLy 269
1132  ACTACTTTAGGCTATCTAACTTGTGTTTGTATTTGCATGTAAACGA 1181
269  sPheAspGIuGIuThrPheleuLysThrleuGIuAspTyrLysCysSers 286
1182  ATTTGACGAGAGAGACTTTTAAAAACATGCAAGATTACAAATTTTCAA 1231
286  erValIleLeuValProThrleuPheAlaIleLeuAsnArgSerGIuLeu 302
1232  GCGTTATCTCTTACCGACTTGTGTCATTTCTTAATGAAGATTA 1281
303  LeuAspLysTyrAspLeuSerAsnleuValGIuIleAlaSerGIuVal 319

```

```

1282  CTCGATTAATATGATTTATCAAAATTTAGTTGAATTTCATCTGCGGAGC 1331
319  aProLeuSerLysGIuIleGIuValAlaValAlaArgArgPheAsnleuP 336
1332  ACCTTTATCTAAAGAAATTTGGTGAAGCTGTGCTAGACGTTTAAATTAC 1381
336  roGIuValArgGIuGIuTyrGIuLeuThrGIuLysThrSerAlaIleIle 352
1382  CCGGTGTTGCTCAAGGCTATGTTTAAACAGAAACCACTTCGCAATTAAT 1431
353  IleThrProGIuGIuLysAspLysProGIuAlaSerGIuLysValValPr 369
1432  ATCAACACCGAGGAGCGATGATTAACCAAGGTCTTGGCAAAATTTGGCC 1481
369  oleuPheLysAlaLysValIleAspLeuAspThrLysLysThrleuGIuP 386
1482  ATTAATTTAAAGCAAAAGTTATTCATCTTGTATCTTAAAAAAAATTTGGGCC 1531
386  roAsnArgArgGIuGIuValCysValLysGIuProMetleuMetLysGIy 402
1532  CGAACAGACGTGAGAGAAAGTTTGTAAAGGCTCTTATGTTAAGATG 1581
403  TyrValAspAsnProGIuAlaArgThrArgGIuIleIleAspGIuGIuLyr 419
1582  TATGTAGATTAATCCAGAGACACAGAAATCATCATGATGAAGAGTTG 1631
419  pleuHisThrGIuAspIleGIuTyrTyrAspGIuGIuLysHisPhePheI 436
1632  GTTGACACACGAGGATATTGGTATTACATGAAAGAAAAACATTTCTTAA 1681
436  leValAspArgLeuLysSerleuIleLysTyrLysGIuTyrGIuValPro 452
1682  TCGTGATGCTTGAAGCTCTTAAATCAAAATCAAGATATCAAGATCA 1731
453  ProAlaGIuLeuGIuSerValleuLeuGIuHisProAsnIlePheAspAl 469
1732  CCGCTGAATTAAGAAATCTGTTCTTTGCAACATCCAAATATTTTGATGC 1781
469  aGIuValAlaGIuValProAspProIleAlaGIuGIuLeuProGIuValAr 486
1782  CGCGGTGCTGGCGTTCCAGATCCATACCTGCTGAGCTTCCGGGAGCTG 1831
486  aIleValleuLysLysGIuLysSerMetThrGIuLysGIuValMetAsp 502
1832  TTGTTCTACTTGAAGAAAGGAAATCATGACGTAAGAAAGATTAATGAT 1881
503  TyrValAlaSerGIuValSerAsnAlaLysArgLeuArgGIuValAr 519
1882  TACGTTGCTAGTCAGATTTCAATGCAAAACGTTGCGGTGGTGTCCG 1931
519  gPheValAspGIuValProLysGIuLeuThrGIuLysIleAspGIuLysA 536
1932  TTTTGTGACGAGAACTAAGAGTCTCACTGCTAAATTTGACGGGTAAAG 1981
536  laIleArgGIuIleleuLysLysProValAlaLysMet 548
1982  CAATTAGAGAAATTAAGTGAAGAAACCAAGTTGCTAAGATG 2019

```

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NM1996.DAT: AAT33851

seq_documentation_block:

ID AAT33851 standard; DNA: 2055 BP.

XX AAT33851;

XX 12-NOV-1996 (first entry)

DE DNA encoding mutant streptavidin/luciferase fusion protein.

XX Mutant; streptavidin; luciferase; fusion protein;

KW Streptomyces avidinii; Luciola lateralis; firefly;

KM recombinant production; industry; ds.

XX

718	GGTGTGCTCCAACTANTGAGATTTCACCTCACTACGGAATGGTTCACAG	767
719	rLeuGlyIleSerLysProThrIleValIlePheSerSerLysGlyLeuA	136
768	TTTGGGATCTCTCAAGCCAACTGTTATTTAGTCTCAAAAAAGGATTTAG	817
136	splyValIleThrValGlnLysThrValIleThrAlaIleLysThrIleVal	152
818	ATTAAGTTATACGTGCATAAAAAAGGTAACGTCTATTAAACCATTTGT	867
153	IleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAspAsnPh	169
868	ATATTGGCACACCAAGGAGATTATAGAGGTTATCAATCATGACCAACTT	917
169	eIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThrY	186
918	TATTTAAAAAACACTCCACAGGTTTCCAAAGGATTCAGTTTAAAACTG	967
186	ALGluValAsnArgLysGlnGlnValIleLeuIleMetAsnSerSerGly	202
968	TAGAAGTTAACCGCAAGAACAAATGCTCTTATTAAGAATCTCTGGGT	1017
203	SerThrArgLysLeuProLysGlyValGlnLeuThrHisGlnLysIleValI	219
1018	TCAACCGGTTGGCCAAAGAGTGTCGCACTTACATGATGAATTTGGTCAC	1067
219	rArgPheSerHisAlaIArgAspProIleTyrGlyAsnGlnValSerPro	236
1068	GGGTTTCTCAGCTACAGTACCAATTTATGGAACCAAGTTTCCACAG	1117
236	lyThrAlaIleLeuThrValValIleProPheHisGlyPheGlyMetPhe	252
1118	GCACGGCATTTTAACTGTAGACCATTCATCATGGTTTGGTATGTTT	1167
253	ThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLeuThrLy	269
1168	ACTACTTAGGCTATCTAACTTGTTGGTTTGGTATGTCATGATTAAAGAA	1217
269	sPheAspGlnGluThrPheLeuLysThrLeuGlnAspTyrLysCysSer	286
1218	ATTTCACGAAAGACTTTTAAAAACCTCAGAGTTTACAAAGTTTCAA	1267
286	eValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGlnLeu	302
1268	GGGTATTCCTGTACCAGACTTGTGTTGCCAATTTCTTAATGAAAGTCAAT	1317
303	IleuAspLysTyrAspLeuSerAsnLeuValGlnIleAlaSerGlyValA	319
1318	CTCGATTAATATGATTTATCAAAATTTACTGTGAATATGTCATGTGGCGGAGC	1367
319	aProLeuSerLysGlnIleGlyGlnAlaValAlaIArgArgPheAsnLeuP	336
1368	ACCTTATCTTAAAGAAATTTGGTGAAGCTGTGCTAGACGTTTAAATTAC	1417
336	roGlyValAlaArgGlnGlyArgGlyLeuThrGlnThrSerAlaIleIle	352
1418	CGGGGTGTGTAAGGCTATGGTTTAAACGAAGAACACTCTGCATTTATT	1467
353	IleThrProGlnGlyAspAspLysProGlyAlaSerGlyLysValValPr	369
1468	ATCACACCGGAAGCGCATATTAACACAGGTGCTTGCCAAAGTTGTGCC	1517
369	oleuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP	386
1518	ATTATTTAAAGCAAAAGTTATGCATCTTGATCTAAAAAAACTTTGGGCC	1567
386	roAsnArgArgGlyGlyValLysValLysGlyProMetLeuMetLysGly	402
1568	CGAACACAGCTGGAGCAAGTTTGTTGTTAAAGGGCTCATGCTTATGAAGCT	1617
403	TyrValAspAsnProGlnLathrArgGlnIleIleAspGlnGlnGlyTyr	419
1618	TATGTATGATATATCCAGAACCAACAAAGAAATCATATGATGAAGAAGTTG	1667

```

419 pleuHisThrGlyAspIleGlyTyrThrAspGluGluLysHisPhePheI 436
      |||||||
1668 GTTGCACACAGGAGATTTGGTATTCGATGAGAAAAACATTTCTTTA 1717
436 ILeValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
      |||||||
1718 TCGTGATGCTTGAAGTCTTTATCAATAAAGATATCAAGTACCA 1767
453 ProAlaGluLeuGluSerValLeuGlnHisProAsnIlePheAspAl 469
      |||||||
1768 CCTGCTGAATTCGAATCTGTTCTTTCACACATCAATAATTTTGATGC 1817
469 aGlyValAlaGlyValProAspProIleAlaGlyLysLeuProGlyAla 486
      |||||||
1818 CGCGGTGGTGGGCTCCAGATCTTAGCTGTGAGCTTCCGGAGCTG 1867
486 aValValLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
      |||||||
1868 TTGTTGTTACTGAAAAAGAAAAATCTATGCTGAAAAAGAAATGATG 1917
503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValAla 519
      |||||||
1918 TACGTTGCTAGTCACTTCAATGCAAAACGTTGCGTGCTGCTCCG 1967
519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLys 536
      |||||||
1968 TTTTGTGACGACGAACTAAAGTCTCACTGCTGTAATTTGACGGTAA 2017
536 ILeIleArgGluIleLeuLysLysProValAlaLysMet 548
      |||||||
2018 CAATTAGAGAAATACTGAAAGAAACGTTGCTAAGATG 2055
seq_name: /SIDS2/gcgdata/geneseq/NA1998.DAT:AAV23580
seq_documentation_block:
ID AAV23580 standard; cDNA to mRNA; 2364 BP.
XX
AC AAV23580;
XX
DT 16-JUL-1998 (first entry)
XX
DE Antibody-Firefly Luciferase fusion protein gene.
XX
KW Firefly Luciferase; antibody-Luciferase fusion protein; ds.
XX
OS Luciola cruciata.
XX
Key Location/Qualifiers
FH 1.2364
FT /*tag= a
FT /transl_except= (pos: 670..672, aa: Glu)
FT /transl_except= (pos: 739..741, aa: Trp)
FT /transl_except= (pos: 1369..1371, aa: Ala)
FT /note= "no stop codon given"
XX
PN JP09187281-A.
XX
PD 22-JUL-1997.
XX
PF 09-JAN-1996; 96JP-0001812.
XX
PR 09-JAN-1996; 96JP-0001812.
XX
(KIKK ) KIKKOMAN CORP.
XX
PA MPI: 1998-275089/25.
XX
DR P-PDB; AAM53882.
XX
PT Antibody-firefly luciferase fused protein - and related products
PT i.e. firefly luciferase fused gene, recombinant DNA and its
PT preparation
XX

```

```

PS Disclosure: Page 13; 17pp; Japanese.
XX
CC This sequence encodes the fusion protein of the invention. The protein is
CC a antibody-firefly luciferase fusion protein, in which an antibody part
CC consisting of a peptide having antibody activity is combined with an
CC enzyme part consisting of firefly luciferase.
XX
SQ Sequence 2364 BP; 697 A; 444 C; 559 G; 664 T; 0 other:

alignment_scores:
      Quality: 2798.00      Length: 545
      Ratio: 5.134      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 99.450

alignment_block:
US-09-581-241-6 x AAV23580 ..
Align seg 1/1 to: AAV23580 from: 1 to: 2364

4 MetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrProI 20
      ::::::::::::::::::::
730 CTCGAGAACGATGAAAAATATGTGTATGCTCTGAACCATTTTACCCCTAT 779
20 eGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetAspArgTyr 37
      |||||||
780 TGAAGGCGATCTGCTGGAGCACATTCGCCAGATATATGATGATGATG 829
37 ILeLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValAspTyr 53
      |||||||
830 CAAACTGTGAGACAAATTCCTTACTAAGCACTACCGGTGCGATAT 879
54 ThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyAlaAlaLeu 70
      |||||||
880 ACGTACGCCGAATCTTACAAAATATCATCTGCTAGGAGAGCTTTAA 929
70 sAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuCysSerGluAsn 87
      |||||||
930 GAATATGTGTTGGTGTGTGATGGAAGATTCGCTATCATGCAATAACT 979
87 yGluGluPhePheIleProValLeuAlaGlyLeuPheIleGlyValGly 103
      |||||||
980 GTGAAGAGTTCTTATCTCTGATTAAGCCGTTTATTTATAGGTGCGCT 1029
104 ValAlaProThrAsnGluIleTyrThrIleuArgGluLeuValHisSer 120
      |||||||
1030 GTGGCTCCACATATGATGAGATTTACACTCTACGTGAATTTGTTACACAGTTT 1079
120 uGlyIleSerLysProThrIleValPheSerSerLysGlyLeuAspL 137
      |||||||
1080 AGCGATCTCTAAGCCACAAATTCATTTAGTTCTTAAAGATTAAGATA 1129
137 yValIlePheThrValGlnLysThrValAlaIleLysThrIleValIle 153
      |||||||
1130 AAGTTATACCTGTACAAAAACGTACCTGCTATTAACCAATTCCTTATA 1179
154 LeuAspSerLysValAspTyrArgLysTyrGlnSerMetAspAspPheI 170
      |||||||
1180 TTGCAGACGAAAGTGAATATAGAGTTATCATCATCGACAACTTTAT 1229
170 eLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThrValG 187
      |||||||
1230 TAAAAAAAACACTCCACCAAGTTTCAAAAGATCAAGTTTAAACTGTAG 1279
187 IuValAsnArgLysGluGlnValAlaLeuIleMetAsnSerSerLysSer 203
      |||||||
1280 AAGTTAACCGCAAGACAAAGTTCCTTTATATGAACTCTTCGGGCTCA 1329
204 ThrGlyLeuProLysGlyValAlaGlnLeuThrHisGluAsnIleValThr 220
      |||||||
1330 ACCGGTTTGCACAAAGTGTGCAACTTACTCATGAATAATTTGTTACGCG 1379
220 gPheSerHisAlaArgAspProIleTyrGlyAsnGlnValSerProGlyT 237

```

```

1380 TTTTCTCAGCGTAGAGATCCATTTATGGAACCAAGTTTCCACGAGCA 1429
237 hrAlaIleuThrValValProPheHisSglYPheGlyMetPheThr 253
1430 CGGCTATTCTTAACTGATAGTACCATTCATCATGTTTGGTATGTTTACT 1479
254 ThrLeuGlyTYrLeuThrCysGlyPheArgIleValMetLeuThrLysPh 270
1480 ACTTTAGGCTATCTAAGTCTGTTGTTTGCATATGTCATGTAAACGAAT 1529
270 eaSPGluGluThrPheLeuLysThrLeuGlnAspTYrLysCysSerSery 287
1530 TGAAGAGAGACTTTTAAACACATGCAAGATTCAAAATGTTCAAGCG 1579
287 alIleuValProThrLeuPheAlaIleuAsnArgSergIleuLeu 303
1580 TTTATCTGTAGCAGACTTGTGTCATCTTAATAGAGTGAATTACTG 1629
304 AspLysTYrAspLeuSerAsnLeuValGluIleAlaSerGlyValAlaPr 320
1630 GATAAATATATTTATCAATTTAGTTGAATTCATCTGGCGAGCACAC 1679
320 OlSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeuProG 337
1680 TTTATCTAAAGAAATGTCACACTGTGCTACACGTTTAAATTACCG 1729
337 LyValArgGlnGlyTYrGlyLeuThrGluThrSerAlaIleIleIle 353
1730 GTTGTCTCAAGGCTATGTTTAAACAGAAACACTGCAATTAATATC 1779
354 ThrProGluLysAspLysProGlyAlaSerGlyLysValAlaProle 370
1780 ACACCGGAGGCGATGATTAACCAAGCTGCTTGCGCAAACTGGCCATT 1829
370 UpheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyPro 387
1830 ATTTAAAGCAAAAGTTTCATCTGTTGATCTAAATAAACTTGGCCCG 1879
387 snArgArgGlyGluValCysValLysGlyProMetLeuMetLysGlyTYr 403
1880 ACACACGTGGAGAGTTGTGTTAAAGGTCCTATGCTTAAGAAAGGTTAT 1929
404 ValAspAsnProGluAlaThrArgGluIleIleAspGluGluTYrPle 420
1930 GTATATATTCAGAGACACAGAGAAATCATATGATGAAGAGTTGTT 1979
420 uHisThrGlyAspIleGlyTYrTYrAspGluGluLysHisPhePheIleV 437
1980 GCACACAGAGAGATATGAGTATGATGATGAAGAAACATTTCTTATCG 2029
437 alaAspArgLeuLysSerLeuIleLysTYrLysGlyTYrGlnValProPro 453
2030 TCGATCGCTTAAAGCTTTAATCAATACAAAGGATATCAAGTCCACCT 2079
454 AlaGluLeuGlnSerValIleuLeuGlnHisProAsnIlePheAspAlaG 470
2080 GCTGATATAGATCTGTTCTTTCACACATTCATATTTTGTATGCCG 2129
470 yValAlaGlyValProAspProIleAlaGlyGluLeuProGlyAlaVal 487
2130 CGTTGCTGGCTGCCAGATCTATAGCTGGTGAAGCTTCCGGAGCTGTG 2179
487 alValLeuLysGlyLysSerMetThrGluLysGluValMetAspTYr 503
2180 TTGTTCTTGAAGAAAGAAATCTATGACTGAAAAGAGATATGATGATTAC 2229
504 ValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValAlaArgPh 520
2230 GTTGTCTATCAAGTTCAATGCAAAAGCTTGGCTGTGTGTCGGTTT 2279
520 eValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysAlaI 537

```

2280 TGTGAGCAGAGTACCTAAAGGTCCTACTGTAAATTTGACGTTAAAGCA 2329

537 leArgGluIleLeuLysProValAlaLysMet 548

2330 TTAGAGAAATCTGAAAGAACAGTGTCTAAGATG 2364

seq_name: /SIDS2/gcdata/geneseq/geneseq/NA1999.DAT:AA25717

seq_documentation_block:

ID AAX25717 standard; cDNA to mRNA; 1656 BP.

AA25717:

21-MAY-1999 (first entry)

DE Firefly luciferase gene #3.

KW Bioluminescent protein; catalytic efficiency; stability; firefly;

KW luciferase; chimeric; luciola cruciata; luciola lateralis;

KW Photinus pyralis; primer; PCR; amplification; ss.

OS Chimeric - Luciola lateralis.

OS Chimeric - Photinus pyralis.

PD W09902697-A1.

PD 21-JAN-1999.

PE 30-JUN-1998; 98WO-JP02936.

PR 08-JUL-1997; 97US-0051917.

PA (KIKK) KIKKOMAN CORP.

PI Hirokawa K, Kajiyama N, Murakami S;

DR WPI; 1999-120898/10.

DR P-PSDB; AAW993366.

XX New bioluminescent protein with improved properties - has greater

PT catalytic efficiency and stability and is obtained by modification

PT of natural precursors

PS Example 3; Page 30-31; 53pp; Japanese.

XX The invention relates to the generation of bioluminescent proteins with

CC improved catalytic efficiency and stability. The proteins are generated

CC by addition, deletion or substitution of amino acids from a precursor

CC protein e.g. firefly luciferases, or by constructing chimeric luciferase

CC proteins from luciferase genes from e.g. luciola cruciata, L. lateralis

CC and Photinus pyralis. This sequence represents an example of a chimeric

CC luciferase gene of the invention.

XX Sequence 1656 BP; 529 A; 283 C; 366 G; 478 T; 0 other;

XX

XX

XX

XX

XX

XX

alignment_scores:

Quality: 2701.00 Length: 543

Ratio: 5.039 Gaps: 0

Percent Similarity: 98.711 Percent Identity: 96.133

alignment_block:

US-09-581-241-6 x AAX25717 ..

Align seg 1/1 to: AAX25717 from: 1 to: 1656

1 MetGluAsnMetGluAsnAspGluAsnIleValTYrGlyProGluProPh 17

1 ATGGAAGAAATGAGAACGATGAAGAAATTTGTATGCTGCTAACCAAT 50

17 eTYrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTYrMet 34

51 TTACCTATTGAAGAGGATCTGCTGAGCACAATTCGCCAATATATATG 100


```

34  sPaRgTYrAlAlaLysLeuGlYAlAlAlaLeuPheThrAsnAlaLeuThrGly 50
|||||
101  ATCGATATGCAAAACTTGGAGCAATGCTTTACTACGACACTTACCGGT 150
51  ValAspTYrThrTYrAlaGluTYrLeuGluLysSerCysCysLeuGlyG 67
|||||
151  GTGATTTATACGTACGCGCAATCTTAGAAAAATCATGCTGTCTAGCA 200
67  uAlaLeuLysAsnTYrGlyLeuValValAspGlyArgIleAlaLeuCys 84
|||||
201  GCGCTTAAAGAAATTAAGTTGGTGTCTTATGAGAAATTCGCTTAGCA 250
84  erGluAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPhe 100
|||||
251  GTGAAACGTGAAGAAATCTTATCTCCGTATAGCCGCTTATATTANA 300
101  GlYValGlyValAlaProThrAsnGluIleTYrThrLeuAsnGlyLeuVal 117
|||||
301  GGTGTCGGGTGCTCCACACTAATGAGATTACACTACGTGAATTTGT 350
117  lHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
|||||
351  TCACAGTTTAGGCATCTCTAACCCAAACATTTGATTTACTTATAAAG 400
134  lYLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
|||||
401  GATYAGATAAAGTTATTAACGTACAAAAAACGCTAACCTGCTATTAAACC 450
151  lIleValIleLeuAspSerLysValAspTYrArgGlyTYrGlnSerMetAs 167
|||||
451  ATTGTTATATGACACGCAAGTGAATTTAGAGTTATCAATTCATGGA 500
167  PaSnPheIleLysLysAsnThrProGlnGlyPheLysGlySerPheL 184
|||||
501  CAACCTTTATTAATAAAACACCTCCACAAGGTTTCAAGAGATCAAGTTTA 550
184  ySThrValGluValAsnArgLysGluGlnValAlaLeuIleMetLysSer 200
|||||
551  AAACGTGAGAAAGTTACCGCAAGAACAAAGTTGCTTTATTAATGAACCT 600
201  SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
|||||
601  TCGGGTTCACACCGGTTTGCACAAAGGTGTCACACTTACTCATGAATAATT 650
217  eValThrArgPheSerHisAlaArgAspProIleTYrGlyAsnGlnValS 234
|||||
651  GGTCACTAGATTTTCTCACGCTAGAGATCCAATTATGTGAACCAAGCTTT 700
234  erProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
|||||
701  CACGAGGACGCGTATTTTAACGTAGTACCATTCATCATGATGGTTTGT 750
251  MetPheThrThrLeuGlyTYrLeuThrCysGlyPheArgIleValMetLe 267
|||||
751  ATGTTTACTACTTAGGCTATCTAAGCTTGTGTTTTCGTATTTGCATCTT 800
267  uThrLysPheAspGluGlnThrPheLeuLysThrLeuGlnAspTYrLysC 284
|||||
801  AAGCAAAATTTGACGAAGAGACTTTTAAAAACACGTCAAGATTACCAAT 850
284  ySSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
|||||
851  GTTCAACGCTTATCTGTACCGCACTTTGTTGCAATTTCTTAATAGAACT 900
301  GluLeuLeuAspLysTYrAspLeuSerAsnLeuValGluIleAlaSerG 317
|||||
901  GAATTTACTCATTAATATGATTTATCAAAATTTAGTGAATTTGCATCGC 950
317  yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
|||||
951  CGAGACACCTTTATCTAAGAAATTTGTAAGCTGTGCTTACGAGCTTTTA 1000

```

```

334  snLeuProGlyValAlaArgGlnGlyTYrGlyLeuThrGlnThrThrSerAla 350
|||||
1001  ATTTACCGGGTGTCTCTCAAGGCTATGTTTAAACAGAAACAACTCTGCA 1050
351  lIleIleIleThrProGluLysAspLysProGlyAlaSerGlyLysVal 367
|||||
1051  ATTATATACACCGGAGGCGATGATTAACCAAGCGTCTTCCGCAAACT 1100
367  lValProLeuPheLysAlaLysValIleAsnLeuAspThrLysLysThrL 384
|||||
1101  TGTGCCATTTATTTAAAGCAAAAGTTATTCATCTTGAATCTTAAAAAATCT 1150
384  euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
|||||
1151  TGGGCCCGCAACAGACCTGAGAAAGTTGTGTAAAGGCTCTTATGCTTAG 1200
401  LysGlyTYrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
|||||
1201  AAAGTTATGTAGATTAATCCAGAACGACACAGAAATCATGATGACAGA 1250
417  uGlyTYrLeuHisThrGlyAspIleGlyTYrTYrAspGluGluLysHisP 434
|||||
1251  AGGTGTTGTCACACAGAGATATGGGTATTACGATGAGAAACAAATTT 1300
434  hePheIleValAspArgLeuLysSerLeuIleLysTYrLysGlyTYrGln 450
|||||
1301  TCTTATCATGTCGATCGTTTGAAGCTTTTATCAATTAACAAAGATATCG 1350
451  ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
|||||
1351  GTGGCCCGCCGCTGAATTTGAAATGATATTGTTTACAAACCCCAACATCT 1400
467  eaSPAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
|||||
1401  CGACGCGGCGCTGCGCAGGCTTCCGACGATGACCCCGGTGAACCTTCCG 1450
484  lYAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
|||||
1451  CCGCCGCTGTGTTTGTGGAGCAGCGAAAGACGATGACGGAAGAAAGATTC 1500
501  MetAspTYrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG 517
|||||
1501  GTGGATTACGTCGCCAGTCAAGTAAACACCGGAAAGTTGCGGAGAG 1550
517  yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
|||||
1551  AGTTGTGTTTGTGACGAAGTAAACGAAAGGTCTTACCGGAACAACTCGACG 1600
534  lYlYAlaIleArgGluLeuLysLys 543
|||||
1601  CAGCAAAATCAGAGAGATCCTCATTAAG 1629
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NM1989.DAT:AA01170
seq_documentation_block:
ID  AA01170 standard: DNA: 1644 BP.
XX
AC  AA01170:
XX
DT  14-JUN-1990 (first entry)
XX
DE  Recombinant luciferase gene.
XX
KW  Luciferase; ATP assay; ss.
XX
OS  Luciola cruciata.
XX
PN  EP301541-A.
XX
XX  01-FEB-1989.
XX
PF  28-JUL-1988; 88BP-0112233.
XX

```


1401 TGATGCTGGTGTTGGCGGGCTCGATGCTCTTACTGCGCAGCTTCAG 1450

484 JVALAVALVALLEULLEULSYLSGLYLSYSERMETHGLULYSLIUAL 500

1451 GAGCGCTTGTGTACTGGGAACGGAAAAAAATATACCGAAAAAGAAATA 1500

501 METASPTYVALALASERGLVALSERASNALALATARGLEARGGLYGI 517

1501 ATGCAATTAGCTGCAGACATCAATTTCAATCAATGCAAACGTTTACGCTG 1550

517 YVALARGPHEVALASPGIUALPROLYSGIULEUHCRLIYSLILEASPG 534

1551 TGTGTGCTTTGTGGATGGAAGTACCTAAAGGCTTACTGGAAGAAATTGACG 1600

534 JLYLSALALAEARGIULEULEULSYLSYSERVALALALYSMET 548

1601 GGAGAGCAATTGAGAAATACCTTAAAGAAACCAAGTTGCTAAAGATG 1644

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:AAQ03801

```
seq_documentation_block;
```

ID	AAQ03801	standard; cDNA; 1644 BP.
vv		

AC AAQ03801;

XX
XX
E

DT 26-FEB-1993 (revised entry)

Recombinant luciferase gene.

KW Luciferase; ATP; E. coli; photon; ds.

OS Synthetic.

PN JP02065780-A.

PD 06-MAR-1990.

PF 01-SEP-1988; 88JP-0216229.

PR 01-SEP-1988; 88JP-0216229.

PA (KIKK) KIKKOMAN CORP.

WPI: 1990-113360/15.

DN F-PSDB; AAK03/88.
XX

PT using *Escherichia* sp. bacteria containing recombinant DNA

PS Claim 2: Page 416 + Fig 3: 20nm. Japanese

Prepn. of Luciferase comprises culturing *Escherichia coli* in media

The recombinant DNA is a vector DNA into which has been inserted a luciferase gene from the culture medium.

Luciferase gene which has the nucleotide sequence below. Luciferase can be produced quickly and efficiently with a few minutes of

CC output, useful in assaying ATP.

50 Sequence 1644 BP; 529 A; 276 C; 338 G; 501 T; 0 other;

alignment_scores:

Quality: 2684.00

Percent Similarity: 99.088 Percent Identity: 93.431

alignment_block:

US-09-581-241-6 x AAQ03801 .

Align seg 1/1 to: AAQ03801 from: 1 to: 1644

1 MetGLuAsMeCtLuAsnAspGluAsnIleValTyrCtIleProGluProPh 17
1 ATGGAAGAAATGGAAGAAACCATGAAATATTGTGAGTGGACCTAAACCGTT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGluLeuArGlySerTyMet 34
34 SPArGtYrTrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 ACCGATATCGAAACAACTGGCGCAATTGCTTTACAATGACGATTACGGT 150
51 ValAspTyrThrTyrAlaGluTyrIleGluLysSerCysCysLeuGly 67
151 GTTGATATTCTTACGCCGCAATACCTGGAGAAATCATGTTGCTAGAAA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
201 ACCTTTGCAAAATTAAGTTGGTTGTTGATGGCAAAATGACGCTTATGCA 250
84 eRGLAsnGlySGLuGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GCGAAACGTGAGAAATTTTTTATTCCGTATATACCGGACCTGTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
301 GGTGAGGTGTTCACCCCAATGATGATTAACATCTTACGTGAACTGCT 350
117 IHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
351 TCACAGTTAGTATCTCTAAACCAACAATTGATTATGATTCTTAAAAAG 400
134 ILeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
451 GCTTGATTAAGATTATTACAGTACGAGAAACAGTACTACTATTAAAC 450
151 ILevalIleLeuAspSerLysValAspTyrArgGlyTyrClnSerMetAs 167
451 ATTGTTATACGTAGATAGCAAACTGATTTATTCAGAGATATCAATGCTGGA 500
167 PAspPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
501 CACCTTTATAAAGAAACACCTCCACCAAGCTTTCACAGATCCAGTTTCA 550
184 ySThValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAGCTGTGAAGTTCACCGTAAAGACAAAGTGCCTCTATTAATGAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnIle 217
601 TCGGGTTTACCGGTTTGCACAAAGCGGTACACACTTACTCACGAAATATC 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
551 AGTCACTGATTTTCTCATGTCAGAGATCCGATTTATGGTATACCAAGTTT 700
234 eRProGlyThrAlaIleLeuThrValValProPheHisIleGlyPheGly 250
701 CACCAAGGACCGCGTGTTTAACTGTCGCTTCCATTCATCATGATGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIe 267
751 ATGTTCACCTACTAGAGGTATTTATTTGTTGTTTTCGTTGTAATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
801 AACCAAAATTCATGAGAAACATTTTAAAAACCTACAAAGATTTAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTACAAAGTATTCTGTACCGACCTGTTTGCATTTCTCAACAAAAAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317

|||||.....
901 GAATTACTCAATAATACGATTTGTCAATTTAGTTCAGATTCATCTCG 950
317 YGlyAlaProLeuSerLysGluIleGlyIuAlaValAlaArgArgPhea 334
|||||.....
951 CGGAGCACCTTTATCAAGAAAGAGTGTGAGCGTTGCTAGACGCTTTA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAla 350
|||||.....
1001 ATCTTCCGGTGTTCGCAAGGTTATGTTTAACAGAAACACATCTGCC 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyValAspGlyLysVa 367
|||||.....
1051 ATTATTTATTCACCGAAGAGACGATTAACCGAGACTTCTGAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
|||||.....
1101 CGTCCCGTGTTTAAAGCAAAAGTATTTGATCTGATACCAAAATCTT 1150
384 euGlyProAsnArgArgGlyValValCysValLysGlyProMetLeuMet 400
|||||.....
1151 TAGGTCCCTAACAGACGTGAGAAAGTTGTGTTAAAGAGACCTATGCTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluI 417
|||||.....
1201 AAAGGTATGTAAATTAATCCGAAAGCAACAAAGAACTTATTCAGAGA 1250
417 uGlyTyrPheuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
|||||.....
1251 AGGTGGCTGCACACCGGATATTGATATTGATGAAAGAAACATTT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
|||||.....
1301 TCTTTATTTGTCGATCGTTGAAGCTTTAATCAATACAAAGATACCAA 1350
451 ValProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
|||||.....
1351 GTACCACTGCGCAATTCGATCCGTTCTTTGCACATCCATCTATCTT 1400
467 eaSpAlaGlyValAlaGlyValProAspProIleAlaGlyLeuLeuProG 484
|||||.....
1401 TGATGCTGTGTGCGCGCTTCCTGATCCTGTAGCTGGCAGCTTCAG 1450
484 lValAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
|||||.....
1451 GAGCCGTTGTGTACTGGAAGCGGAAAAAATATGACCGAAAAAGAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgLysI 517
|||||.....
1501 ATGATTTATGTTGCAAGTCAGATTCAATGCAAAACGTTACGTGTGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
|||||.....
1551 TCTTGTGTTTGTGATGAAGTACCTAAAGTCTTACTGAAAAAATTTGACG 1600
534 lYlYsAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
|||||.....
1601 GCAGAGCAATTAGAGAAATCTTAAGAAACGATTGCTTAGATG 1644

Db 481 elpavvllkkgkmsktekewdyvaasvsnakrlrgvrtvdevpkqltqkldgkairrl 540
 Oy 541 LKKPVAKM 548
 Db 541 lkkpvakm 548

RESULT 3

AAW04212 ID AAW04212 standard; Protein: 548 AA.

AC AAW04212;

DT 12-NOV-1996 (first entry)

DE Luciola lateralis luciferase.

KM Mutant; streptavidin; luciferase; fusion protein;

KM Streptomyces avidinii; Luciola lateralis; firefly;

OS recombinant production; Industry.

PN Luciola lateralis.

PD JP07289264-A.

XX 07-NOV-1995.

XX 27-APR-1994; 94JP-0090275.

XX 27-APR-1994; 94JP-0090275.

XX (KIKK) KIKKOMAN CORP.

DR WPI; 1996-015269/02.

XX New mutant streptavidin (SA) gene - fused to a firefly luciferase

PT gene, for the recombinant prepn. of a SA-FL fused protein

PS Disclosure; Page 10; 12pp; Japanese.

XX The present sequence is the L. lateralis luciferase, which can be

CC used in the construction of a streptavidin/luciferase fusion

CC protein, comprising a mutant Streptomyces avidinii streptavidin

CC and the L. lateralis (firefly) luciferase. The fusion protein

CC can be prepd. by inserting the recombinant DNA encoding it into a

CC Escherichia species microbe, culturing the transformed microbe in a

CC medium and collecting the fusion protein from the culture. The

CC fusion protein can be used in industry.

XX Sequence 548 AA:

Query Match 99.9%; Score 2819; DB 17; Length 548;

Best Local Similarity 99.8%; Pred. No. 2.5e-235;

Matches 547; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MEMMENDENIVGPEPEPIEESGAGAOIRKYMRYAKIGATFATNALTGVDYTYAEYLE 60

Db 1 memmendenivgpepfiyeegsagaglrkymryakigatfatafnalgydytyaeyle 60

Oy 61 KSCCLGALKNYGLVVDGRIALCSENCDEFFIYVLGLFVGAVAPTNELTYLRELVHSL 120

Db 61 kscclgealnknyglvvdgrialcseenceeffivlaglfigvavaptneltylrelvhs 120

Oy 121 GISKPTIVSSKKGLDKVITVOKTVAITVTLVLSKVDYRKGQSDNFTKKMTPOGFGK 180

Db 121 giskptivsskkgldkvitvoktvtaitvtlvtlsvdvyrgysmdnftkkmtppgfkx 180

Oy 241 TVVFNHGFGEFTTGLYLTGCFRIYVMTLTKFDEETPLKTDYDKCSSVILVPLFAILNRS 300

Db 241 tvvpfnhgfgefttglyltcgfrivymtltkfdeetfkltdydkcssvllvplfai 300

Oy 301 ELIDKYDLSNLYEIASGAPLSKEIGEAARPNLPVNRQYGLTETTSATIIITTEBGDK 360

Db 301 elldkydlsnlveyiasgapslskeigearpnlpvnrqyglteetsailltebgddk 360

Oy 361 PGASGVVPLFKAVIDLDLTKLGPNRREVCVKGMIMKGVNDPEATREIIEEGWL 420

Db 361 pgasgvvplfkavidldltkclgpnrrevcvkgpmlmkgyvndpeatreiideegwl 420

Oy 421 HTGDIGYDEEKEHFTVDRLKSLIKYKGYOVPADESVLLOHPNIFDAGVAVDPPIAG 480

Db 421 htgdigydeekhfvtvdrllkslikykygvpaelesvllqhnifdagvavdpplag 480

Oy 481 ELPGAVVVLKKGKSMTEKEVMDYVASQVSAKALRGVRFVDEVPKGLKIDGAKAREI 540

Db 481 elpgavvvllkkgkmsktekewdyvaasvsnakrlrgvrtvdevpkqltqkldgkairrl 540

Oy 541 LKKPVAKM 548

Db 541 lkkpvakm 548

XX RESULT 4

XX AAM53884

XX ID AAM53884 standard; Protein: 548 AA.

XX AAM53884;

XX 16-JUL-1998 (first entry)

XX Antibody-Firefly Luciferase fusion protein.

XX Firefly Luciferase; antibody-luciferase fusion protein.

XX Synthetic.

XX Luciola cruciata.

XX JP09187281-A.

XX 22-JUL-1997.

XX 09-JAN-1996; 96JP-0001812.

XX 09-JAN-1996; 96JP-0001812.

XX (KIKK) KIKKOMAN CORP.

XX WPI; 1998-275089/25.

XX Antibody-firefly luciferase fused protein - and related products

XX i.e. firefly luciferase fused gene, recombinant DNA and its

XX preparation

XX Disclosure; Page 9-10; 17pp; Japanese.

XX This sequence is a fusion protein of the invention. The protein is

XX a antibody-firefly luciferase fusion protein, in which an antibody part

XX consisting of a peptide having antibody activity is combined with an

XX enzyme part consisting of firefly luciferase.

XX Sequence 548 AA:

Query Match 99.9%; Score 2819; DB 19; Length 548;

Best Local Similarity 99.8%; Pred. No. 2.5e-235;

Matches 547; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MEMMENDENIVGPEPEPIEESGAGAOIRKYMRYAKIGATFATNALTGVDYTYAEYLE 60

Db 1 memmendenivgpepfiyeegsagaglrkymryakigatfatafnalgydytyaeyle 60

QY 61 KSCCIGELKNTGLVVDGRIALCSENCEEFFIPVLAGLFIGVAPNTEIYTLRELVSL 120
 DB 61 KSCCIGELKNTGLVVDGRIALCSENCEEFFIPVLAGLFIGVAPNTEIYTLRELVSL 120
 QY 121 GISKPTIVSSKKGIDKVTYTKTATKTYIILDSKVDYRGYSMDNFIRKNTPOQFKG 180
 DB 121 GISKPTIVSSKKGIDKVTYTKTATKTYIILDSKVDYRGYSMDNFIRKNTPOQFKG 180
 QY 181 SSFKTEVNRKEQVALINMSSGSTGLPKGYOLTHENLVTRFSHARDPIYGNQVSPGTAI 240
 DB 181 SSFKTEVNRKEQVALINMSSGSTGLPKGYOLTHENLVTRFSHARDPIYGNQVSPGTAI 240
 QY 241 TVPPEHGEFMTTGLYITCGFRIVMLTKDEEFLKTLQDYKCSSVILVPTLFAILNRS 300
 DB 241 TVPPEHGEFMTTGLYITCGFRIVMLTKDEEFLKTLQDYKCSSVILVPTLFAILNRS 300
 QY 301 ELDDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGYGLTETSATITTPBEDDK 360
 DB 301 ELDDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGYGLTETSATITTPBEDDK 360
 QY 361 PGASGKVPPLFKAVIDLTKTTLGPNRGEVCVKGPMKMGVYDNPDEATREIIDEEGWL 420
 DB 361 PGASGKVPPLFKAVIDLTKTTLGPNRGEVCVKGPMKMGVYDNPDEATREIIDEEGWL 420
 QY 421 HTGDIgyDEEKHFFIYDRLKSLIKYGYOPPALESVILQHPNTRDAGVAPPPDIAG 480
 DB 421 HTGDIgyDEEKHFFIYDRLKSLIKYGYOPPALESVILQHPNTRDAGVAPPPDIAG 480
 QY 481 ELPGAIVVLLKKGSMTEKEVMDVYASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIRREI 540
 DB 481 ELPGAIVVLLKKGSMTEKEVMDVYASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIRREI 540
 QY 541 LKKPVAKM 548
 DB 541 LKKPVAKM 548

RESULT 5

AAM99367
 ID AAM99367 standard; Protein; 548 AA.

AC AAM99367;

DT 21-MAY-1999 (first entry)

DE Mutant firefly luciferase protein #4.

KM Bioluminescent protein; catalytic efficiency; stability; firefly;

KW luciferase; chimeric; Luciola cruciata; Luciola lateralis;

OS Photinus pyralis; primer; PCR; amplification.

OS Luciola lateralis.

OS Synthetic.

PN WO9902697-A1.

PD 21-JAN-1999.

PF 30-JUN-1998; 98WO-JP02936.

PR 08-JUL-1997; 97US-0051917.

PA (KIKK) KIKKOMAN CORP.

PI Hirokawa K, Kajiyama N, Murekami S;

DR WPI; 1999-120898/10.

DR N-PSDB; AAX25718.

PT New bioluminescent protein with improved properties - has greater
 PT catalytic efficiency and stability and is obtained by modification
 PT of natural precursors

XX Example 5; Page 36-39; 53pp; Japanese.
 PS The invention relates to the generation of bioluminescent proteins with
 CC improved catalytic efficiency and stability. The proteins are generated
 CC by addition, deletion or substitution of amino acids from a precursor
 CC protein e.g. firefly luciferase, or by constructing chimeric luciferase
 CC proteins from luciferase genes from e.g. Luciola cruciata, L. lateralis
 CC and Photinus pyralis. This sequence represents an example of a mutated
 CC luciferase protein of the invention.
 XX

Sequence 548 AA:

Query Match. 99.8%; Score 2817; DB 20; Length 548;
 Best Local Similarity 99.8%; Pred. No. 3.8e-235;
 Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEMMENDENIVYGPPEFYPIEGSAGAOILRYMDRBAKILGATAFNMLTGVDYTYAEYLE 60
 DB 1 memmendenivyppefypiegsaqlrkymdryakigalatafnaltgydytyaeyle 60
 QY 61 KSCCIGELKNTGLVVDGRIALCSENCEEFFIPVLAGLFIGVAPNTEIYTLRELVSL 120
 DB 61 KSCCIGELKNTGLVVDGRIALCSENCEEFFIPVLAGLFIGVAPNTEIYTLRELVSL 120
 QY 121 GISKPTIVSSKKGIDKVTYTKTATKTYIILDSKVDYRGYSMDNFIRKNTPOQFKG 180
 DB 121 GISKPTIVSSKKGIDKVTYTKTATKTYIILDSKVDYRGYSMDNFIRKNTPOQFKG 180
 QY 181 SSFKTEVNRKEQVALINMSSGSTGLPKGYOLTHENLVTRFSHARDPIYGNQVSPGTAI 240
 DB 181 SSFKTEVNRKEQVALINMSSGSTGLPKGYOLTHENLVTRFSHARDPIYGNQVSPGTAI 240
 QY 241 TVPPEHGEFMTTGLYITCGFRIVMLTKDEEFLKTLQDYKCSSVILVPTLFAILNRS 300
 DB 241 TVPPEHGEFMTTGLYITCGFRIVMLTKDEEFLKTLQDYKCSSVILVPTLFAILNRS 300
 QY 301 ELDDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGYGLTETSATITTPBEDDK 360
 DB 301 ELDDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGYGLTETSATITTPBEDDK 360
 QY 361 PGASGKVPPLFKAVIDLTKTTLGPNRGEVCVKGPMKMGVYDNPDEATREIIDEEGWL 420
 DB 361 PGASGKVPPLFKAVIDLTKTTLGPNRGEVCVKGPMKMGVYDNPDEATREIIDEEGWL 420
 QY 421 HTGDIgyDEEKHFFIYDRLKSLIKYGYOPPALESVILQHPNTRDAGVAPPPDIAG 480
 DB 421 HTGDIgyDEEKHFFIYDRLKSLIKYGYOPPALESVILQHPNTRDAGVAPPPDIAG 480
 QY 481 ELPGAIVVLLKKGSMTEKEVMDVYASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIRREI 540
 DB 481 ELPGAIVVLLKKGSMTEKEVMDVYASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIRREI 540
 QY 541 LKKPVAKM 548
 DB 541 LKKPVAKM 548

RESULT 6

AAR30803
 ID AAR30803 standard; Protein; 548 AA.

AC AAR30803;

DT 24-MAY-1993 (first entry)

DE Thermostable luciferase (wild-type).

DE Firefly; Luciola cruciata; GENJ1; Luciola lateralis; HEIKE;

KW mutant; assay.

XX Luciola lateralis.

XX Key Location/Qualifiers
 FH Misc-difference 217
 FT /note= "amino acid replaced by hydrophobic amino
 FT acid, esp. Ile, Leu or Val, for heat stable
 FT mutant"
 EP524448-A.
 XX
 XX 27-JAN-1993.
 PD
 XX 26-JUN-1992; 92EP-0110808.
 PE
 XX 27-JUN-1991; 91JP-0157117.
 PR 29-NOV-1991; 91JP-0317064.
 PR 22-MAY-1992; 92JP-0131057.
 XX
 PA (KIKK) KIKKOMAN CORP.
 PI Ellich N, Naoki K;
 DR WPI: 1993-028553/04.
 DR N-PSDB: AAO34745.
 XX
 PT New DNA sequence of a thermostable luciferase - has the aminoacid
 PT at position-217 of Luciola cruciata or lateralis luciferase
 PT replaced by a hydrophobic aminoacid
 PS
 XX Claim 10-12; Page 30 + 25-29; 33pp; English.
 CC A thermostable luciferase of a firefly, in which an amino acid
 CC at the 217 position or an amino acid at the position equiv. to the
 CC 217 position of the luciferase of Luciola cruciata (GENUI firefly)
 CC of Luciola lateralis (HEIKE firefly) is replaced by a hydrophobic
 CC amino acids, esp. Ile, Leu or Val, is claimed. The mutant luciferase
 CC is identical in properties to the wild-type luciferase except that
 CC it is stable when heated to high temps., e.g. 50 degrees C.
 CC The luciferase can be used in assays for e.g. ATP.
 SQ Sequence 548 AA:

Query Match 99.7%; Score 2814; DB 14; Length 548;
 Best Local Similarity 99.6%; Pred. No. 6.9e-235;
 Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MENMENDENIYGPPEFYIEEGSAGQIRKYMRYAKIGATFATNALTGVDYTAETYLE 60
 DB 1 memmendeniygppefyieegsagaqlrkymdryakigatfatnaltgvdtyaeyle 60
 QY 61 KSCCLDEALKNGLVVDGRIALCSENCFFIPVLAGLFGVAPNTNITLRELVAHL 120
 DB 61 ksccldealknglvvdgrialcseenceeffipvlaglfgvavpntnityclrelvahl 120
 QY 121 GIKPTIVSSKGGDLKVTYVOKTVAIKTIVILDSKVDYRGYSMDNEIKNTPGFKG 180
 DB 121 gikptivsskggdlkvtvoktvaiktivildskvdrygysmdneikntpgfkfg 180
 QY 181 SSKRYEVRNKEEOVALIMNSGSGTGLPKYQVLTHENLVYRFSNARPIYGNQVSPETALL 240
 DB 181 sskryevrnkeeovalimnsgsgtgtpkyqvlthenavtrfsharpiygnqvspetall 240
 QY 241 TVVPRPHNGFEMFTLGYLNGCFRIWMLTRPDEETFLTKDODYCSSLVITPPLFALLINS 300
 DB 241 tvvprphngfemftlgylnngcfriwmltrpdeetfltkdodycsslvitpplfallins 300
 QY 301 ELIDKYLDSNLVIEIASGAPLSEIGEAARPNLPGVRGYSGLTETSAIIITPBGDK 360
 DB 301 elldkylslnlviesagaplskeigearpnlpgvrqysglteetsaiiitpbgdk 360
 QY 361 PGASGKVVPLFKAKVIDDIDKTKTLGPRRGEVCKGPMILKKGVDNDEARETIIDEGM 420
 DB 361 pgasgkvvplfkakvididkttkltgprngevcvkgpmlkgyvndnpearetiidegwl 420

QY 421 HTGDICGYDEEKHFIVDRLSLIKKGYPAPLESVLLQHPNIFDAGVAGVDPPIAG 480
 DB 421 htgdigydydeekhfivdrllslikkygypaalesvllqhpnidagvavgvpdpdiag 480
 QY 481 ELPGAVVVLKKGSKMTEKEVMYVASOVSNARLRGVRFEVDEPKLTGKIDKAIREI 540
 DB 481 elpgavvvlekksmtekevmvyasovsnarlrgrvrfdevpklgtgkidakairei 540
 QY 541 LKRPVAKM 548
 DB 541 lkrpvakm 548

RESULT 7
 AAR3710
 ID AAR3710 standard; Protein; 548 AA.
 XX
 XX AAR3710;
 AC
 XX
 XX 24-MAY-1993 (first entry)
 DT
 XX
 DE Thermostable luciferase (T217L,L,V).
 XX
 XX Firefly; Luciola cruciata; GENUI; Luciola lateralis; HEIKE;
 KM mutant; assay.
 XX
 OS Luciola lateralis.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 217
 FT /label= ILE, LEU, VAL
 FT
 XX
 XX EP524448-A.
 PN
 XX 27-JAN-1993.
 PD
 XX 26-JUN-1992; 92EP-0110808.
 PE
 XX 27-JUN-1991; 91JP-0157117.
 PR 29-NOV-1991; 91JP-0317064.
 PR 22-MAY-1992; 92JP-0131057.
 XX
 XX (KIKK) KIKKOMAN CORP.
 PA
 PI Ellich N, Naoki K;
 DR WPI: 1993-028553/04.
 DR
 XX
 XX New DNA sequence of a thermostable luciferase - has the aminoacid
 PT at position-217 of Luciola cruciata or lateralis luciferase
 PT replaced by a hydrophobic aminoacid
 PS
 XX Claim 10-12; Page 30 + 25-29; 33pp; English.
 CC A thermostable luciferase of a firefly, in which an amino acid
 CC at the 217 position or an amino acid at the position equiv. to the
 CC 217 position of the luciferase of Luciola cruciata (GENUI firefly)
 CC of Luciola lateralis (HEIKE firefly) is replaced by a hydrophobic
 CC amino acids, esp. Ile, Leu or Val, is claimed. The mutant luciferase
 CC is identical in properties to the wild-type luciferase except that
 CC it is stable when heated to high temps., e.g. 50 degrees C.
 CC The luciferase can be used in assays for e.g. ATP.
 SQ Sequence 548 AA:

Query Match 99.7%; Score 2814; DB 14; Length 548;
 Best Local Similarity 99.6%; Pred. No. 6.9e-235;
 Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MENMENDENIYGPPEFYIEEGSAGQIRKYMRYAKIGATFATNALTGVDYTAETYLE 60
 DB 1 memmendeniygppefyieegsagaqlrkymdryakigatfatnaltgvdtyaeyle 60

Db 1 memendeniyypepfypleegsagaqlrkymdryaklgalaifnaltygvdyaeyle 60
 QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPVLAGFTIGVAPRTNEITRLVHSL 120
 Db 61 kscclgealknyglvvdgrialscenceeffipvlagftigvaptneitrlvhs 120
 QY 121 GISKPTIVFSSKKGLDKVITYOKTVAIKTVILDSKVDRGYOSMDNFIRKNTPOGFGK 180
 Db 121 giskptivfsskglkdvityoktvaiktviltdskvdrgyosmdnfirkntpogfgk 180
 QY 181 SSFKTEVNRKEQVALIMNSGSTGLPKGVOLTHENLVTRFSHARDPIYGNQVSPGTAIL 240
 Db 181 ssfktevnrkeqvalimnssgstglpkgvqlthenxvtrfshardpiygnvspgtail 240
 QY 241 TYVPEHHGFMFTTGLYLCGFRIVMLTKFDETFELKTLQDYKCSSVILVPTLFAILNRS 300
 Db 241 tyvpfhngfmitclglylfcgfrivmltkfdeetfkltdqykcssvllvptlfaillrs 300
 QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVRQGYGLTETTSAILITPEGDDK 360
 Db 301 elldkydlslnlveiasgapslskeigearrnlpgvrtgyglteetsailitpegddk 360
 QY 361 PGASGKVVPLFKAKVIDLDRKKTLPNRGECVCKGPMIMKGYVDNDEATREIIDEEGWL 420
 Db 361 pgasgkvvplfkakvidldrktlgnprngevcvkgpmlmkgyvdpnpeatreideegwl 420
 QY 421 HTGDIGYDEEKHFYVDRLKSLIKRYGVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 Db 421 htgdigydeekhfivdrllkslikrygvppaesvllqhpnlfdagvagvdpdpiag 480
 QY 481 ELPGAVVVLKKGSMTEKEVMDYVASOVSNARKRLRGVRVDVPGKLTGKIGKAIKREI 540
 Db 481 elpgavvvlekysmtekevmdyvasovsnarkrlrgvrvdvpgkltgkldgkaiirel 540
 QY 541 LKKPVAKM 548
 Db 541 lkkpvakm 548

RESULT 8
 AAM12395
 ID AAM12395 standard; Protein: 636 AA.
 AC AAM12395;
 DT 20-MAY-1997 (first entry)
 DE Firefly luciferase conjugated with 87 aa biotin peptide at 3' end.
 KW Fusion protein: firefly; luciola lateralis; luciferase; biotinylation;
 KW wild type; E.coli; bioluminescence assay; quantification; ligand;
 KW receptor.
 OS Synthetic.
 FH Key
 FT Region Location/Qualifiers
 FT Region 1..547
 FT Region /note="firefly luciferase portion of fusion protein"
 FT Region 548..549
 FT Region /note="encoded by linker sequence used to construct
 FT Region fusion gene"
 FT Region 549..636
 FT Region /note="biotin peptide portion of fusion protein"
 PN JP08308578-A.
 PD 26-NOV-1996.
 XX 24-APR-1995; 95JP-0098857.
 XX 14-MAR-1995; 95JP-0054625.
 PR 27-JUL-1994; 94JP-0193798.
 XX

PA (KIKK) KIKKOMAN CORP.
 XX
 DR WPI: 1997-059697/06.
 DR N-PSDB; AAT63269.
 XX
 PT Fusion protein comprising firefly luciferase and biotinylated
 PT peptide - useful in a bio-luminescent analytical method for
 PT quantifying ligands
 XX
 PS Claim 2; Page 12-13; 13pp; Japanese.
 XX
 CC This is the amino acid sequence of a novel fusion protein which comprises
 CC the firefly (Luciola lateralis) luciferase protein and an 87 amino acid
 CC biotinylated peptide designated the E.coli biotin carboxy-carrier protein
 CC (BCCP-87). The fusion gene was generated by firstly amplifying the
 CC sequence encoding the E.coli BCCP-87 from the E.coli genome and inserting
 CC the resultant sequence into the 3' end of the firefly luciferase gene in
 CC plasmid pHLf230. The biotin peptide sequence replaces the C-terminal Met
 CC residue of the wild type luciferase sequence. The plasmid was then
 CC introduced into E.coli JM101 for production of the fusion protein. The
 CC luciferase ligands which may modulate binding of luciferase to its
 CC receptor.
 CC
 CC Sequence 636 AA;
 XX
 SQ

Query Match 99.7%; Score 2814; DB 18; Length 636;
 Best Local Similarity 99.8%; Pred. No. 8, 6e-235;
 Matches 546; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEMNENDENIYVPEPFYPIEGSAGALRKYMDRYAKLGAIFTNALNGVDVYAEYLE 60
 Db 1 memendeniyypepfypleegsagaqlrkymdryaklgalaifnalngvdyaeyle 60
 QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPVLAGFTIGVAPRTNEITRLVHSL 120
 Db 61 kscclgealknyglvvdgrialscenceeffipvlagftigvaptneitrlvhs 120
 QY 121 GISKPTIVFSSKKGLDKVITYOKTVAIKTVILDSKVDRGYOSMDNFIRKNTPOGFGK 180
 Db 121 giskptivfsskglkdvityoktvaiktviltdskvdrgyosmdnfirkntpogfgk 180
 QY 181 SSFKTEVNRKEQVALIMNSGSTGLPKGVOLTHENLVTRFSHARDPIYGNQVSPGTAIL 240
 Db 181 ssfktevnrkeqvalimnssgstglpkgvqlthenlvtrfshardpiygnvspgtail 240
 QY 241 TYVPEHHGFMFTTGLYLCGFRIVMLTKFDETFELKTLQDYKCSSVILVPTLFAILNRS 300
 Db 241 tyvpfhngfmitclglylfcgfrivmltkfdeetfkltdqykcssvllvptlfaillrs 300
 QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVRQGYGLTETTSAILITPEGDDK 360
 Db 301 elldkydlslnlveiasgapslskeigearrnlpgvrtgyglteetsailitpegddk 360
 QY 361 PGASGKVVPLFKAKVIDLDRKKTLPNRGECVCKGPMIMKGYVDNDEATREIIDEEGWL 420
 Db 361 pgasgkvvplfkakvidldrktlgnprngevcvkgpmlmkgyvdpnpeatreideegwl 420
 QY 421 HTGDIGYDEEKHFYVDRLKSLIKRYGVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 Db 421 htgdigydeekhfivdrllkslikrygvppaesvllqhpnlfdagvagvdpdpiag 480
 QY 481 ELPGAVVVLKKGSMTEKEVMDYVASOVSNARKRLRGVRVDVPGKLTGKIGKAIKREI 540
 Db 481 elpgavvvlekysmtekevmdyvasovsnarkrlrgvrvdvpgkltgkldgkaiirel 540
 QY 541 LKKPVAKM 547
 Db 541 lkkpvak 547

RESULT 9

AA048855	AA048855 standard; Protein; 640 AA.
ID	AA048855 standard; Protein; 640 AA.
XX	
XX	AA048855;
AC	
XX	
DT	10-SEP-1998 (first entry)
XX	
DE	Luciferase-lystostaphin fused protein.
XX	
KW	Luciferase-lystostaphin fused protein; chimeric; bioluminescence
XX	
OS	Chimeric - Staphylococcus simulans.
OS	Chimeric - Luciola lateralis.
XX	
FH	Key
FH	Location/Qualifiers
FT	Domain
FT	1..549
FT	/note= "Luciferase protein"
FT	550..640
FT	/note= "lystostaphin protein"

PN	JPI0150991-A.
XX	
PD	09-JUN-1998.
XX	
PF	25-NOV-1996; 96JP-0328042.
XX	
PR	25-NOV-1996; 96JP-0328042.
XX	
PA	(KIKK) KIKKOMAN CORP.
XX	
XX	WPI; 1996-379994/33.
DR	N-PSDB; AAV32467.
DR	
XX	
PT	New luciferase-lysoctaphin fused protein - useful for
PT	bioluminescence analysis
XX	
PS	Claim 4; Pages 7-9; 10pp; Japanese.

The invention claims for a luciferase-lyso-staphin fused gene which encodes a fused protein in which a peptide part consisting of luciferase, from *Lucifera lateralis*, is connected to a peptide part consisting of lyso-staphin from *Staphylococcus simulans*. The method described by the invention can be used to prepare a luciferase-lyso-staphin fused protein efficiently. The fused protein is useful for bioluminescence analysis.

SQ Sequence 640 AA

Query Match	99.78%	Score 2814	DB 19	Length 640
Best Local Similarity	99.88%	Pred. No. 8.7e-235		
Matches 546: Conservative	1	Mismatches 0	Indels 0	Gaps 0

Oy	1	MENMEDENIVGPEPPYPIIEEGSAGOLRKMYRPAKLAIAFTNALGVDTVAEYLE	60
Dd	1	menmendeniyygpeppiyieegsaagqlktymdryakigafistnalgvdytlyeayle	60
Oy	61	KSCCIGEAALKNTGLVWDGRIALCSBCEBEFFPYLACLGFGVAVAPNEIYTIRELVHSL	120
Dd	61	kscclgealknxyilvwdgrialcsenceeeffipylaglfigyvapntneilytlrelvhsl	120
Oy	121	GISKRPITVSSKKGDKITVQKFTYALTAKTITVLIDSKVDYDRGYOSMNFKNKTPOGFGK	180
Dd	121	glskrpitvssakkgyldkvltvygktvtaktltilidskvdvydrygysmndfknktpbgfgk	180
Oy	181	SSEFTEVEVNRKDOVALIMNSSGTCLPKGVOLTIENTLYTRFSHARDPITYGNVSPTGALL	240
Dd	181	sseftevevnrtkeqvalimmsgstclprkyvalitheanlytrfsahardpygnqvspgtall	240
Oy	241	TVPVRPHHGSMFTTLGUYLCGCRFIYMALKPDEEFFLTLDQYKSSVIATPFTFALLNS	300
Dd	241	tvpvrphhgsmfttlygltcgrfiymalkpdeeffltkldqykssviatpftfallns	300

Oy	301	ELDKYDLSNLEVASGAPLSKETIGNAVARFNLBCVRCGYLLETETSAIIITPBGDDK	360
Db	301	elldkydlsnlevasgaplsketigavarfnlbpvrgyylteltsailitpbgdk	360
Oy	361	PGASGKVVPLTKAVVLDLDTKKTGLPNNRREVCVKSGPMLKKGVDNDEATREIIDEQWL	420
Db	361	pgasgkvvpltkavvldltdtkktlgnprtrgevcvksgpmlkkyvdnpeatreiideeqwl	420
Oy	421	HTGDIGYDEKHEFIYDRKSLTKYRGYQVPAELESYLLQHNPNRPDACAQVPPIDG	480
Db	421	htgdigydeekhfitydrksltkyrgyqvpaeelesyllqhnlpdagsvqppidg	480
Oy	481	ELPGAVVLLKRGKSGTEKEVMDVYASOVSNNAKRLRGVRFVDEVPKGLTGKIDGKAIREI	540
Db	481	elpgavvllkrgksktekevmdvyasovsnakrlrgvrfvdevpkgltgkldgkairei	540
Oy	541	LKKPVAK	547
Db	541	lkkpvak	547

Db 541 1kkpvak 547

```

RESULT 10
AAR03731
ID AAR03731 standard; protein; 548 AA

```

AA
AC AAR03731;

DT 22-JUL-1990 (first entry)

Luciola lateralis recombinant luciferase.

XX
KW Luciferase; enzyme; ATP; plasmid pHLf7; firefly.

XX
XX
Lucio Jatoris

XX
PN EP353464-A.

07-FEB-1990.
PD

AA 30-JUN-1989; 89EP-0111958.
PF

XX
DB 33-DEC-1988. 88JP-0162402

PR 01-JUL-1988; 88JP-0162402.

PA (KIKK) KIKKOMAN CORP.

xx
PI
Tatsumi H, Kajiyama N, Nakano E;

XX
XX
WBT: 1990-038240/06
DB

DR N-PSDB; AAQ03257.

PT New gene encoding luciferase -

PT from *Escherichia coli* genus cells.

XX
DC
Claim 3: page 31: 43nn: English

XX

contained in plasmid pHLf7.

AA 548 AA;
SO Sequence

Query Match	99.68;	Score 2811;	DB 11;	Length 548;
Best Local Similarity	99.58;	Pred. No. 1.2e-234;		
Matches 545; Conservative	2;	Mismatches 1;	Indels 0;	Gaps 0;

Qy 1 MEMEMDENIVYGPPEFFPIIEGSGAOLRKIMRYAKLGIATNNLTGVDTIYAETLE 60
| | | | |
Db 1 memmemdenivygpepfiiiegsgagaqlrkymdtyaklgiafntalgtvdylyeayle 60
| | | | |
Qy 61 KSCCLTEALKNGLVVDGRIALCSNCEBF1PVIAGLFICGVGAAPINIEITTLRELVHSL 120
| | | | |

```

Db 61 kscclgealnknyjlvdgrialcsenceeffipvlaglflygvavptneilylrelvhs1 120
QY 121 GISKPTIVFSSKRGKLDKVTYVOKTVAIKTIVILDSKVYGYOSMNFKKNTPOGFKG 180
Db 121 giskptivfsskrgkldkvtvqktvtaiktvlildskvdyrgysmdnflknlpqgfkq 180
QY 181 SSFTVEVNRKEQVALIMNSSGSLGPKGVOLTHENLVTRFSHARDPIYGNVSPGTAIL 240
Db 181 ssftvevnrkeqvalimnssgstlgpkgyvqlthenavtrfshardpiygnvsgptail 240
QY 241 TVVPHHGFCEMFTTIGYTCGFRIVMLTKPDEEFELTKLDYKCSSYILVPTLFAILNRS 300
Db 241 tvvpfhgfmfttlylgtcgrivmltkfdeefeltkldykcossyilvptlfailnrs 300
QY 301 ELDDKYLNLVETASGAPLSKEIGAVARRFNLPGVRGQYGLTETTSAILITPEGDDK 360
Db 301 elldkylnlveiasgapslskeigavarrrfnlpgvrqygltettsailitpegddk 360
QY 361 PGASGVVPLFKAVIDLDTKTIGPVRGECVCKGPMLMKGYVDNEATREITIDEGWL 420
Db 361 pgasgvvplfkavidldtktlgpnrrgevcvkqgpmimkgyvdneatreitideegwl 420
QY 421 HTGDIGYDEEKHEFFIVDRLSLIKYGQVPAELSEVLLQHPNIFDAGVAGVPDPIAG 480
Db 421 htgdigydeekheffivdrllkslikygyvppaelsevlilqhpnlfdagvavpdpia 480
QY 481 ELGAVVVLKKGKSMTEKEMDYVASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
Db 481 elgavvvllkkgksmtkemydyvasqvsnakrlrgvrfvdelpkgltgkldgkaiirei 540
QY 541 LKKPVAKM 548
Db 541 lkkpvakm 548

RESULT 11
AAW12394
ID AAW12394 standard; Protein; 568 AA.
AC
XX AAW12394;
XX
XX 20-MAY-1997 (first entry)
DE Firefly luciferase conjugated with 23 aa biotin peptide at N-terminus.
XX
XX Fusion protein: firefly: luciola lateralis; luciferase; biotinylation;
XX wild type; E.coli; bioluminescence assay; quantification; ligand;
XX receptor.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX FT 1..24
XX FT /note= "biotin peptide"
XX FT 25..568
XX FT /note= "firefly luciferase"
XX
XX JP08308578-A.
XX
XX 26-NOV-1996.
XX
XX 24-APR-1995; 95JP-0038857.
XX
XX 14-MAR-1995; 95JP-0054625.
XX 27-JUL-1994; 94JP-0193798.
XX
XX (KIKK ) KIKKOMAN CORP.
XX
XX WPI: 1997-059697/06.
XX DR N-PSDB; AAT63268.
XX
XX Fusion protein comprising firefly luciferase and biotinylated
XX peptide - useful in a bio-luminescent analytical method for

```

```

PT quantifying ligands
XX
XX Claim 2: Page 10-11; 13pp; Japanese.
PS
XX
XX This is the amino acid sequence of a novel fusion protein comprising
CC the firefly (luciola lateralis) luciferase protein and a 23 amino acid
CC biotinylated peptide designated biotin peptide #84. The gene was
CC generated by firstly annealing the sequence encoding peptide #84
CC (AAT63270) and its complementary sequence, then inserting the resultant
CC double stranded sequence into the 5' end of the luciferase gene in
CC plasmid pHLf108. The biotin peptide sequence replaces the first 4 amino
CC acids of the wild type luciferase sequence. The plasmid was then
CC introduced into E.coli JM101 for production of the fusion protein. The
CC novel protein can be used in bioluminescence assays to quantify
CC luciferase ligands which may modulate binding of luciferase to its
CC receptor.
XX
XX Sequence 568 AA:
SQ

```

```

Query Match 99.5%; Score 2801; DB 18; Length 568;
Best Local Similarity 99.5%; Pred. No. 9,7e-234;
Matches 543; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 NMENDENTVYGEPEPYPTIEGSAQAQLRKYMRYAKLGAIAFTNALGVDTVAEYLEKS 62
Db 23 slendenlvypgepyptiegsagaqlrkymdryaklgaiaftnalgvdtyaeyleks 82
QY 63 CCLGALNKYGLVNDGRALALCSENCEFFIPVLAGLFGVADPTNETYLRELVHS1GI 122
Db 83 cclgealnkylvndgrialalcsenceeffipvlaglflygvavptneilylrelvhs1gi 142
QY 123 SKPTIVFSSKRGKLDKVTYVOKTVAIKTIVILDSKVYRGYOSMNFKKNTPOGFKGSS 162
Db 143 skptivfsskrgkldkvtvqktvtaiktvlildskvdyrgysmdnflknlpqgfkss 202
QY 163 FKTVEVNRKEQVALIMNSSGSLGPKGVOLTHENLVTRFSHARDPIYGNVSPGTAILTY 242
Db 203 fktvevnrkeqvalimnssgstlgpkgyvqlthenavtrfshardpiygnvsgptailty 262
QY 243 VPFHGFCEMFTTIGYTCGFRIVMLTKPDEEFELTKLDYKCSSYILVPTLFAILNRS 302
Db 263 vpfhgfcmfttlylgtcgrivmltkfdeefeltkldykcossyilvptlfailnrsel 322
QY 303 LDKYDLNLVETASGAPLSKEIGAVARRFNLPGVRGQYGLTETTSAILITPEGDDKPG 362
Db 323 ldkydlnlveiasgapslskeigavarrrfnlpgvrqygltettsailitpegddkpg 382
QY 363 ASGVVPLFKAVIDLDTKTIGPVRGECVCKGPMLMKGYVDNEATREITIDEGWLHT 422
Db 383 asgvvplfkavidldtktlgpnrrgevcvkqgpmimkgyvdneatreitideegwlht 442
QY 423 GDIGYDEEKHEFFIVDRLSLIKYGQVPAELSEVLLQHPNIFDAGVAGVPDPIAGEL 482
Db 443 gdigydeekheffivdrllkslikygyvppaelsevlilqhpnlfdagvavpdpia 502
QY 483 PGAVVVLKKGKSMTEKEMDYVASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIREIL 542
Db 503 pgavvvllkkgksmtkemydyvasqvsnakrlrgvrfvdelpkgltgkldgkaiireil 562
QY 543 KPVAKM 548
Db 563 kpvakm 568

RESULT 12
AAW53883
ID AAW53883 standard; Protein; 568 AA.
AC
XX AAW53883;
XX
XX 16-JUL-1998 (first entry)
XX

```

DE Antibody-Firefly Luciferase fusion protein.
 XX Firefly luciferase; antibody-luciferase fusion protein.
 XX Synthetic.
 OS Luciola cruciata.
 XX JP09187281-A.
 PN 22-JUL-1997.
 XX 09-JAN-1996; 96JP-0001812.
 XX 09-JAN-1996; 96JP-0001812.
 PR 09-JAN-1996; 96JP-0001812.
 XX (KIKK) KIKKOMAN CORP.
 PA WPI; 1998-275089/25.
 DR N-PSDB; AAV23595.
 XX Antibody-firefly luciferase fused protein - and related products
 PT i.e. firefly luciferase fused gene, recombinant DNA and its
 PT preparation
 XX Disclosure; Page 11-13; 17pp; Japanese.
 PS This sequence is a fusion protein of the invention. The protein is
 CC an antibody-firefly luciferase fusion protein, in which an antibody part
 CC consisting of a peptide having antibody activity is combined with an
 CC enzyme part consisting of firefly luciferase.
 XX
 XX Sequence 568 AA:

Query Match 99.2%; Score 2801; DB 19; Length 568;
 Best Local Similarity 99.5%; Pred. No. 9.7e-234;
 Matches 543; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 NMENDENIVGPPPEYPIEGSAGALRKYMRYAKLGAIAFTNALTGVDYTYAEYLEKS 62
 Db 23 slendeniivgpepfybiegsagqlrkymdyaklgatalfnaltgydytyaeyleks 82
 QY 63 CCGEALKNYGLVVDGRIALCSENCEEFIPVLAGLFTIGGVAPTNEIYTLRELTVHSLGI 122
 Db 83 cclgealknyglvvdgrialscenseefipvlaglftigvaptneiytlreltvhslgi 142
 QY 123 SKPTIVSSKKGLDKVITVOKTYTAIKTIVILDSKVDYRGYSMDNFIRKNTQGFGRSS 182
 Db 143 skptivsskkglkvltvqktvtaiktivildskvdyrgysmdnfirkntpqgfgkgs 202
 QY 183 EKTVEVNRKEQVALIMSSSGSTGLPKVOULTHEMLVTRFSHADPIYGNQVSPGTATLTV 242
 Db 203 fktvevnrkeqvalimsssgstglpkvoulthemlvtrfshadpiygnqvsgptaltlv 262
 QY 243 VPFHGGFMFTTLYTCGFRIVMLTFKDEBETFLKTLQDYKCSSVILVPTLFAILNSEL 302
 Db 263 vpfhggfmfttlytgcgfrivmltkfdeetflkldqdykcssvilvptlfaalnrsel 322
 QY 303 LDKYDLSMLVEIASGAPLSEKEIGENARRNLPGVQGVLTETTSALITTPGDDKPG 362
 Db 362 ldkydlsmelveiasgaplskeigeavarfnlpgvrgyyltettsaliltpegddkpg 382
 QY 383 ASGVKVPLEFKAKVIDLDRKKTGLPNNRGEVCVKGPMLKAGVNDPEATREITDEGWLHT 422
 Db 422 asgvkvpelfkavldlrrkktglpnnrgevcvkgpmlkagvndpeatreitideegwlht 442
 QY 442 GDIGYDEEKEHFIIVDRLSLTKIKKGYOVPAELESVLLQHPNITFDAGVAPPIAGEL 482
 Db 482 gdigydeekhefiivdrslsltkikkgvovpaelesvllqhpnitfdagvappiagel 502
 QY 502 PGAVVVLTKKSGMTEKEVNDVYASOVNAKRLRGVRFVEDEPKGLTKIDGAKIRELIK 542
 Db 542 pgavvvltkksgmtekevndvyasovnakrlrgvrfvedepkgltkidgakirelik 562

QY 543 KPYAKM 548
 Db 563 kvakm 568

RESULT 13
 AAM04208
 ID AAM04208 standard; Protein; 673 AA.
 XX
 XX AAM04208;
 AC 12-NOV-1996 (first entry)
 XX
 XX Streptavidin/luciferase fusion protein.
 DE Streptavidin; luciferase; fusion protein;
 XX Streptomyces avidinii; Luciola lateralis; firefly;
 KW recombinant production; industry.
 KW
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH 1..128
 FT Protein /label= streptavidin
 FT 130..673
 FT Protein /label= luciferase
 XX
 XX JP07289264-A.
 PN 07-NOV-1995.
 XX 27-APR-1994; 94JP-0090275.
 XX 27-APR-1994; 94JP-0090275.
 PR 27-APR-1994; 94JP-0090275.
 XX (KIKK) KIKKOMAN CORP.
 PA WPI; 1996-015269/02.
 DR N-PSDB; AAT33850.
 XX New mutant streptavidin (SA) gene - fused to a firefly luciferase
 PT gene, for the recombinant prepn. of a SA-FL fused protein
 XX
 XX Example 1; Pages 6-7; 12pp; Japanese.
 PS The present sequence is a streptavidin/luciferase fusion
 CC protein, comprising the Streptomyces avidinii streptavidin gene
 CC and the Luciola lateralis (firefly) luciferase gene. The fusion
 CC protein can be prep'd by inserting the recombinant DNA encoding
 CC it into a Escherichia species microbe, culturing the transformed
 CC microbe in a medium and collecting the fusion protein from the
 CC culture. The fusion protein can be used in industry.
 XX
 XX Sequence 673 AA:

Query Match 99.2%; Score 2801; DB 17; Length 673;
 Best Local Similarity 99.5%; Pred. No. 1.2e-233;
 Matches 543; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 NMENDENIVGPPPEYPIEGSAGALRKYMRYAKLGAIAFTNALTGVDYTYAEYLEKS 62
 Db 128 slendeniivgpepfybiegsagqlrkymdyaklgatalfnaltgydytyaeyleks 187
 QY 63 CCGEALKNYGLVVDGRIALCSENCEEFIPVLAGLFTIGGVAPTNEIYTLRELTVHSLGI 122
 Db 188 cclgealknyglvvdgrialscenseefipvlaglftigvaptneiytlreltvhslgi 247
 QY 123 SKPTIVSSKKGLDKVITVOKTYTAIKTIVILDSKVDYRGYSMDNFIRKNTQGFGRSS 182
 Db 248 skptivsskkglkvltvqktvtaiktivildskvdyrgysmdnfirkntpqgfgkgs 307

QY 163 FKTVEVNRKEQVALIMNSSGSTGLPKGVOLTHENLVTRFSHARDPIYGNVSPGTAITLV 242
 Db 308 fktvevnrkeqvalimnssgstglpkgyqlthenlvtrfshardpiyngvsgptalltv 367
 QY 243 VPFHFGMFTTGLYLCGFRIVMLTKRDEETFLKTLQDYKCSVILVPLFAILNSEL 302
 Db 368 vpfhfgmfttglylctgfivmltkrdeetflklqdykcsvillvplfailnssel 427
 QY 303 LDKYDLSNLVETIASGAPLSEKEIGEAARFNLPGVRQGYLTETTSAILITPEGDDKPG 362
 Db 428 ldkydlnslveiasgapslekeigearfnlpgvrqgylltettailltpegdckpg 487
 QY 363 ASGVVPLFKKSKMTEKEVMDYVASQVSNARKRLGVRFPVDEVRKGLTGKIDGKAIREILK 422
 Db 488 asgvvplfkakvldltdkktlgnrrgevcvkgpmlkgyvdpnpeatrelldeegwllht 547
 QY 423 GDIGYDEEKHFIVDLKSLIKKGYVPPAELESVILQHPNIFDGVAGVDPDPIAGEL 482
 Db 548 gdigydeekhfivdlkslikkyvppaelsevlilqhpnidagvagvdpdpiaagel 607
 QY 483 PGAVVVLKKGKSMTEKEVMDYVASQVSNARKRLGVRFPVDEVRKGLTGKIDGKAIREILK 542
 Db 608 pgavvvlekkgksmtekevmdyvasqvsnakrltgyvrfvdevpkyltgdgkatrellk 667
 QY 543 KPYAKM 548
 Db 668 kpyakm 673

RESULT 14

AAW04209
 ID AAW04209 standard; Protein; 685 AA.

AC AAW04209;

DT 12-NOV-1996 (first entry)

DE Mutant streptavidin/luciferase fusion protein.

KW Mutant; streptavidin; luciferase; fusion protein;
 KW Streptomyces avidinii; Luciola lateralis; firefly;
 KW recombinant production; industry.

OS Synthetic.

PH Key Location/Qualifiers

FT Protein 1..140
 FT Protein /label= mutant streptavidin 142..685

FT Peptide /label= luciferase 1..17

FT /label= mutant N-terminal peptide

PN JP07289264-A.

XX 07-NOV-1995.

XX 27-APR-1994; 94JP-0090275.

XX 27-APR-1994; 94JP-0090275.

XX (RIKK) KIRKOMAN CORP.

XX WPI; 1996-015269/02.

XX N-PSDB; AAT33851.

XX New mutant streptavidin (SA) gene - fused to a firefly luciferase

XX gene, for the recombinant prepn. of a SA-FL fused protein

XX Claim 4; Pages 8-9; 12pp; Japanese.

XX The present sequence is a mutant streptavidin/luciferase fusion

XX protein, comprising a mutant Streptomyces avidinii streptavidin and

CC the Luciola lateralis (firefly) luciferase. The fusion protein can
 CC be prepd. by inserting the recombinant DNA encoding it into a
 CC Escherichia species microbe, culturing the transformed microbe in a
 CC medium and collecting the fusion protein from the culture. The
 CC fusion protein can be used in industry.

CC Sequence 685 AA:

Query Match 99.2%; Score 2801; DB 17; Length 685;
 Best Local Similarity 99.5%; Pred. No. 1.3e-233;

Matches 543; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 NNDENIYVGEPEPYPIEESGAGQLRKYMDRVAKLAIFTNALTGVDVTAIEYLS 62
 Db 140 slendenlyvgepefpylieesagagllkymdryaklaiaiftnaltgvdtlyeayls 199
 QY 63 CCLGEALKNYGLVVDGRALALCSECEFFIYVLAGLFIGVGAAPNEIYTLRELVHSLGI 122
 Db 200 cclgealknvglvvdgralalcsenceeffiyvlaglfigvgaapneiytlrelvhsigl 259
 QY 123 SKPTIVFSSKKGLDKVITVOKTVAIKTIVILDSKVDRGYQSDMNFKKMTPGCFKSS 182
 Db 260 skptivfsskkgldkvitvoktvaiktivildskvdrgyqsdmnfkkmtpgcfkss 319
 QY 183 FKTVEVNRKEQVALIMNSSGSTGLPKGVOLTHENLVTRFSHARDPIYGNVSPGTAITLV 242
 Db 320 fktvevnrkeqvalimnssgstglpkgyqlthenlvtrfshardpiyngvsgptalltv 379
 QY 243 VPFHFGMFTTGLYLCGFRIVMLTKRDEETFLKTLQDYKCSVILVPLFAILNSEL 302
 Db 380 vpfhfgmfttglylctgfivmltkrdeetflklqdykcsvillvplfailnssel 439
 QY 303 LDKYDLSNLVETIASGAPLSEKEIGEAARFNLPGVRQGYLTETTSAILITPEGDDKPG 362
 Db 440 ldkydlnslveiasgapslekeigearfnlpgvrqgylltettailltpegdckpg 499
 QY 363 ASGVVPLFKKSKMTEKEVMDYVASQVSNARKRLGVRFPVDEVRKGLTGKIDGKAIREILK 422
 Db 500 asgvvplfkakvldltdkktlgnrrgevcvkgpmlkgyvdpnpeatrelldeegwllht 559
 QY 423 GDIGYDEEKHFIVDLKSLIKKGYVPPAELESVILQHPNIFDGVAGVDPDPIAGEL 482
 Db 560 gdigydeekhfivdlkslikkyvppaelsevlilqhpnidagvagvdpdpiaagel 619
 QY 483 PGAVVVLKKGKSMTEKEVMDYVASQVSNARKRLGVRFPVDEVRKGLTGKIDGKAIREILK 542
 Db 620 pgavvvlekkgksmtekevmdyvasqvsnakrltgyvrfvdevpkyltgdgkatrellk 679
 QY 543 KPYAKM 548
 Db 680 kpyakm 685

RESULT 15

AAW53882
 ID AAW53882 standard; Protein; 788 AA.

AC AAW53882;

DT 16-JUL-1998 (first entry)

DE Antibody-Firefly luciferase fusion protein.

XX Firefly luciferase; antibody-luciferase fusion protein.

XX Synthetic.

XX Luciola cruciata.

XX Key Location/Qualifiers

FT Misc-difference 224 /note= "encoded by CAG"

FT Misc-difference 247

FT /note= "encoded by GAT"
FT Misc-difference 457
FT /note= "encoded by TTG"

PN JP09187281-A.

PD 22-JUL-1997.

PF 09-JAN-1996; 96JP-0001812.

PR 09-JAN-1996; 96JP-0001812.

PA (KIKK) KIKKOMAN CORP.

DR WPI: 1998-275089/25.

DR N-PSDB; AAV233580.

XX Antibody-firefly luciferase fused protein - and related products
PT i.e. firefly luciferase fused gene, recombinant DNA and its
PT preparation

PS Claim 5; Page 14-16; 17pp; Japanese.

CC This sequence is the fusion protein of the invention. The protein is
CC a antibody-firefly luciferase fusion protein, in which an antibody part
CC consisting of a peptide having antibody activity is combined with an
CC enzyme part consisting of firefly luciferase.

XX Sequence 788 AA:

Query Match 98.7%; Score 2785; DB 19; Length 788;

Best Local Similarity 99.3%; Pred. No. 3.8e-232;

Matches 541; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 MENDENIYGPPEPPIEESGASAGQIRKYMDRYAKLGAIAFTNALGVDTYAEYLEKSC 63

DB :|||||

QY 244 lenweniygpefypiegsagaglrkymdryaklgafafnaltgvdyaeyleksc 303

DB CIGEAALKNGIYVVDGRIALCSNCEEFFIPVLAGLFGIGVAPTEIYTRRLVHSLGIS 123

DB 304 clgealknqylvvdgrlalcsenceeffipvlagllfgvavaptnelyclrelvmslgis 363

QY 124 KPTIVSSKKGIDKVTYVQKTYTAIKTIYILDSKVDYRGQSMDFIKKPTQGFKSSSF 183

DB :|||||

DB 364 kptivsskkgldkvtvqktaiktlvldskvdrygysmdnfiknptqgfksssf 423

QY 184 KTVENRRKEQVALIMNSSGSTGLPRGVQLTHENLVTRFSHARDPIYGNQVSPGTALITVY 243

DB :|||||

DB 424 ktvevnrkkeqvalimnssgstglprkvqlthenavtrfshardpiygnqvspgtalittv 483

QY 244 PRHNGGMMFTTGLYLCGFRIVMLTFRDEBFLTKLQDYKCSSVTLVPLFAILNRSELL 303

DB :|||||

DB 484 prhnggmfttglylclgrivmltkfdeeftikldqdykcssvllvplclfaillnrsell 543

QY 304 DKYDLNLIWEIASGAPLSEKEIGEAVARRENLPQVROGYGLTETTSAILITPEGDDKPGA 363

DB :|||||

DB 544 dkydlnliveiasgagplsekeigearrenlpqvrqgyglteetsailltpgeddkpga 603

QY 364 SGRVVPLEFAKVIDLDTKTKTLGPNRRGEVCKGPMIMKGYVDNPEATFREI IDEEGWLATG 423

DB :|||||

DB 604 sgkvvpelfkakvidldtktktlgpnrrgevcvqpmimkgyvdnpeatreiideegwlhtg 663

QY 424 DIGYDEEKHEFTVDRKLKSLIKKGYOVPAALESVLLQHPNIFDAGVAGVDPDTAGELP 483

DB :|||||

DB 664 digydeekheftvdrklkslikkygvpaaesvllqhpnlfdagvagvpdpdtagelp 723

QY 484 GAVVVLKKGSMTEKEMVYASQVSNARLRGRVFEVPEPKGLTGKIDGKAIREILKK 543

DB :|||||

DB 724 gavvvleksgsmtekevmdyvasqvsnakrlrgvrfvdevpkgltgkldgkairleilk 783

QY 544 PVAKM 548

DB :|||||

DB 784 pvakm 788

Search completed: November 29, 2001, 01:53:12
Job time: 3252 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:53:12 ; Search time 46.79 Seconds
(without alignments)
867.539 Million cell updates/sec

Title: US-09-581-241-6

Perfect score: 2823
Sequence: 1 MENNENDENTIVYGEPEPPYI.....TGKIDGKAIREILKKPYAKM 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq.1101.*
1: /SID2/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2823	100.0	548	AA1980	L. lateralis lucif
2	2821	99.9	548	AA1981	L. lateralis lucif
3	2817	99.8	548	AA1982	Antibody-Firefly 1
4	2817	99.8	548	AA1983	Mutant Firefly Luc
5	2815	99.7	548	AA1984	Thermotable Lucif
6	2814	99.7	548	AA1985	Thermotable Lucif
7	2814	99.7	548	AA1986	Firefly Luciferase
8	2812	99.6	548	AA1987	Luciferase-Lysosoma
9	2812	99.6	548	AA1988	Luciola lateralis
10	2811	99.6	548	AA1989	Firefly Luciferase
11	2799	99.1	548	AA1990	

12	2799	99.1	548	AA1991	Antibody-Firefly 1
13	2799	99.1	548	AA1992	Streptavidin/Lucif
14	2799	99.1	548	AA1993	Mutant streptavidin
15	2785	98.7	788	AA1994	Antibody-Firefly 1
16	2701	95.7	552	AA1995	Chimeric firefly 1
17	2684	95.1	548	AA1996	Recombinant lucifer
18	2684	95.1	548	AA1997	Luciferase, Synth
19	2684	95.1	548	AA1998	Modified Luciola c
20	2681	95.0	548	AA1999	L. cruciata lucifer
21	2680	94.9	548	AA2000	Thermotable lucif
22	2680	94.9	548	AA2001	Thermotable lucif
23	2678	94.9	548	AA2002	Luciferase mutant
24	2678	94.9	548	AA2003	L. cruciata lucife
25	2678	94.9	548	AA2004	L. cruciata lucife
26	2677	94.8	548	AA2005	Luciferase, Lucio
27	2676	94.8	548	AA2006	Luciferase mutant
28	2676	94.8	548	AA2007	L. cruciata lucife
29	2674	94.7	548	AA2008	Luciferase mutant
30	2671	94.6	548	AA2009	Luciferase mutant
31	2671	94.6	548	AA2010	Chimeric firefly 1
32	2669	94.5	548	AA2011	Chimeric firefly 1
33	2590	91.7	552	AA2012	Luciferase of Hota
34	2394	84.8	552	AA2013	Modified firefly 1
35	2343.5	83.0	548	AA2014	Plant signal trans
36	1956.5	69.3	550	AA2015	Modified firefly 1
37	1956.5	69.3	550	AA2016	lux::npt-II fusion
38	1947.5	69.0	551	AA2017	Firefly Luciferase
39	1946.5	68.9	550	AA2018	P. pyralis pteLuc
40	1945.5	68.9	550	AA2019	Amino acid sequenc
41	1945.5	68.9	550	AA2020	Firefly Luciferase
42	1945.5	68.8	550	AA2021	P. pyralis variant
43	1942.5	68.7	550	AA2022	
44	1939.5	68.7	550	AA2023	
45	1939.5	68.7	550	AA2024	

ALIGNMENTS

RESULT 1
AA1980 standard; Protein: 548 AA.
AC AAY22183;
DT 09-SEP-1999 (first entry)
DE L. lateralis luciferase protein sequence.
XX Luciferase; surfactant-tolerant; firefly; intracellular ATP assay.
XX Luciola lateralis.
XX OS
XX PN WO9933997-A1.
XX PD 08-JUL-1999.
XX PF 24-DEC-1998; 98WO-JP05864.
XX PR 26-DEC-1997; 97JP-0361022.
XX PA (KIRK) KIRKOMAN CORP.
XX PI Hattori N, Murakami S;
XX DR WPI: 1999-419109/35.
XX DR N-PSDB; AAX84396.
XX PT Modified Luciferase tolerant to surfactants and useful for assay of
XX PF intracellular ATP
XX PS Claim 5; Page 43-48; 56pp; English.
XX XX

Db 481 elpgavvvllkqkmsktekmdyvasqvsnakrlrgrvrvfdevpkglgkldgkxalrei 540
 QY 541 LKKPVAKM 548
 Db 541 LKKPVAKM 548

RESULT 3

AAW04212
 ID AAW04212 standard: Protein: 548 AA.

AC AAW04212;

DT 12-NOV-1996 (first entry)

XX Luciola lateralis luciferase.

XX Mutant; streptavidin; luciferase; fusion protein;

XX Streptomyces avidinii; Luciola lateralis; firefly;

XX recombinant production; industry.

XX Luciola lateralis.

XX JP07289264-A.

XX 07-NOV-1995.

XX 27-APR-1994; 94JP-0090275.

XX 27-APR-1994; 94JP-0090275.

XX (KIKK) KIKKOMAN CORP.

XX WPI; 1996-015269/02.

XX New mutant streptavidin (SA) gene - fused to a firefly luciferase

XX gene, for the recombinant prep. of a SA-FL fused protein

XX Disclosure; Page 10; 12pp; Japanese.

XX The present sequence is the L. lateralis luciferase, which can be

XX used in the construction of a streptavidin/luciferase fusion

XX protein, comprising a mutant streptomyces avidinii streptavidin

XX and the L. lateralis (firefly) luciferase. The fusion protein

XX can be prep. by inserting the recombinant DNA encoding it into a

XX Escherichia species microbe, culturing the transformed microbe in a

XX medium and collecting the fusion protein from the culture. The

XX fusion protein can be used in industry.

XX Sequence 548 AA:

XX

XX

XX

XX

XX

Query Match 99.8%; Score 2817; DB 17; Length 548;
 Best Local Similarity 99.6%; Pred. No. 3.4e-235;
 Matches 546; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNMNDENIYVGPPEPTPIEGSAGALRKYMRYAKLGAIAFTNALTGVDTYAEYLE 60
 Db 1 MNNMNDENIYVGPPEPTPIEGSAGALRKYMRYAKLGAIAFTNALTGVDTYAEYLE 60
 QY 61 KSCCLGEALKNYGLVVDKRIALCSENCEFFIPLVLALEIGVVAFTNEIYTLRELHSL 120
 Db 61 KSCCLGEALKNYGLVVDKRIALCSENCEFFIPLVLALEIGVVAFTNEIYTLRELHSL 120
 QY 121 GISKPTIVSSKKGIDKRYTVOKTYATKYTYILDSKYDVRGYSMDNFIRKNTPOGFRK 180
 Db 121 GISKPTIVSSKKGIDKRYTVOKTYATKYTYILDSKYDVRGYSMDNFIRKNTPOGFRK 180
 QY 181 SSPKTYEVNRKQVALIMNSGSGTGLPKGVOLTHENITVRFESHADPTIYGNQVSPCTAIL 240
 Db 181 SSPKTYEVNRKQVALIMNSGSGTGLPKGVOLTHENITVRFESHADPTIYGNQVSPCTAIL 240

QY 241 TVPVEHHGFMETTLGLYTCGRFIVMLTKFDEETFLKTLQDYKCSSVILVPTLFAILNRS 300
 Db 241 tvpvhngfmgfttlgyltgcgrfivmltkfdeetflkltldgkcssvllvptlfailnrs 300
 QY 301 ELDDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGIGLTETTSAILITPEBDDK 360
 Db 301 elldkysdlnslveiasgapslskeigearfnlpgvrgygltetssailitpegddk 360
 QY 361 PGASGVVPLFKAVYIDDTKKTGGPNRGEVCYKGMALMGVYDNPEATREIIDEEGWL 420
 Db 361 pgasgvvplfkavvldtktggnrgevcykgpmlmgvydnpeatreideegwl 420
 QY 421 HTGDIGYDEEKHEFTVDRLSLIRKYGQVPPAELESVLLQHPNIFDAGVAGVPPIAG 480
 Db 421 htgdigydeekheftvdrslslikygyvppaelsevlqhpnifdaagvayvpdpia 480
 QY 481 ELPGAIVVVLKKSMTKEKEMDYNVAVSQVSNKRLRGVRFVDEVPKGLTGKIDGKALREI 540
 Db 481 elpgavvvllkqkmsktekmdyvasqvsnakrlrgrvrvfdevpkglgkldgkxalrei 540
 QY 541 LKKPVAKM 548
 Db 541 LKKPVAKM 548

RESULT 4

AAW53884
 ID AAW53884 standard: Protein: 548 AA.

AC AAW53884;

DT 16-JUL-1998 (first entry)

XX Antibody-firefly luciferase fusion protein.

XX Firefly luciferase; antibody-luciferase fusion protein.

XX Synthetic.

XX Luciola cruciata.

XX JP09187281-A.

XX 22-JUL-1997.

XX 09-JAN-1996; 96JP-0001812.

XX 09-JAN-1996; 96JP-0001812.

XX (KIKK) KIKKOMAN CORP.

XX WPI; 1996-275089/25.

XX Antibody-firefly luciferase fused protein - and related products

XX i.e. firefly luciferase fused gene, recombinant DNA and its

XX preparation

XX Disclosure; Page 9-10; 17pp; Japanese.

XX This sequence is a fusion protein of the invention. The protein is

XX an antibody-firefly luciferase fusion protein, in which an antibody part

XX consisting of a peptide having antibody activity is combined with an

XX enzyme part consisting of firefly luciferase.

XX Sequence 548 AA:

XX

XX

XX

Query Match 99.8%; Score 2817; DB 19; Length 548;
 Best Local Similarity 99.6%; Pred. No. 3.4e-235;
 Matches 546; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNMNDENIYVGPPEPTPIEGSAGALRKYMRYAKLGAIAFTNALTGVDTYAEYLE 60
 Db 1 MNNMNDENIYVGPPEPTPIEGSAGALRKYMRYAKLGAIAFTNALTGVDTYAEYLE 60

```

QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFIPVLAGLFGVAVAPNTEITREIVHSL 120
DB 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFIPVLAGLFGVAVAPNTEITREIVHSL 120
QY 121 GISKPTIVSSSKGLDKVITVQKTAIATITVILDSKVDRGYOSMDNFKKNTPOGFKG 180
DB 121 GISKPTIVSSSKGLDKVITVQKTAIATITVILDSKVDRGYOSMDNFKKNTPOGFKG 180
QY 181 SSFKTEVNRKEQVALIMNSSGSTGLPKGVOLTHENIVRFSHARDPIRYGNOSPGATL 240
DB 181 SSFKTEVNRKEQVALIMNSSGSTGLPKGVOLTHENIVRFSHARDPIRYGNOSPGATL 240
QY 241 TVPFFHGHGFMFTTGLYLCGFRIVMLTKFDEETPLKTLQDYKCSSVILPTLFAILNRS 300
DB 241 TVPFFHGHGFMFTTGLYLCGFRIVMLTKFDEETPLKTLQDYKCSSVILPTLFAILNRS 300
QY 301 ELLDKYDLNIVEIASGAPLSKEIGEAVARRNLCGVNRQVGLTETTSATITTPGDDK 360
DB 301 ELLDKYDLNIVEIASGAPLSKEIGEAVARRNLCGVNRQVGLTETTSATITTPGDDK 360
QY 361 PGASGVVPLFKAKVIDDITDKTKTLGPNRRGEVCVKGPMIMKGVNDPPEATREITDEEGL 420
DB 361 PGASGVVPLFKAKVIDDITDKTKTLGPNRRGEVCVKGPMIMKGVNDPPEATREITDEEGL 420
QY 421 HNGDIGYDEEKNHFIYDRKSLIKYKGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
DB 421 HNGDIGYDEEKNHFIYDRKSLIKYKGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
QY 481 ELPGAUVVLKKGSMTEKEVMDYVASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIKREI 540
DB 481 ELPGAUVVLKKGSMTEKEVMDYVASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIKREI 540
QY 541 LKKPVAKM 548
DB 541 LKKPVAKM 548

RESULT 5
AAM9367
ID AAM9367 standard; Protein: 548 AA.
AC AAM9367;
XX 21-MAY-1999 (first entry)
DE Mutant firefly luciferase protein #4.
XX Bioluminescent protein; catalytic efficiency; stability; firefly;
KW luciferase; chimeric; luciola cruciata; luciola lateralis;
KM photinus pyralis; primer; PCR; amplification.
XX Luciola lateralis.
OS Synthetic.
XX MO9902697-A1.
XX 21-JAN-1999.
XX 30-JUN-1998; 98MO-JP02936.
XX 08-JUL-1997; 97US-0051917.
XX (KIKK ) KIKKOMAN CORP.
XX Hirokawa K, Kajiyama N, Murakami S;
XX WPI: 1999-120898/10.
XX N-PSDB; AAX25718.
XX New bioluminescent protein with improved properties - has greater
PT catalytic efficiency and stability and is obtained by modification
PT of natural precursors

```

```

XX Example 5; Page 36-39; 53pp; Japanese.
PS The invention relates to the generation of bioluminescent proteins with
XX improved catalytic efficiency and stability. The proteins are generated
CC by addition, deletion or substitution of amino acids from a precursor
CC protein e.g. firefly luciferase, or by constructing chimeric luciferase
CC proteins from luciferase genes from e.g. luciola cruciata, L. lateralis
CC and photinus pyralis. This sequence represents an example of a mutated
CC luciferase protein of the invention.
XX
SQ Sequence 548 AA;
Query Match 99.7%; Score 2815; DB 20; Length 548;
Best Local Similarity 99.6%; Pred. No. 5e-235;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MENMENDENIVYGPPEPYPIEESAGAOQLRKYMDRAKIGATAFTALGVDTYVAEYLE 60
DB 1 menmendenivyppepypieesagaglrkymdryakigataftalgvdytyaeyle 60
QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFIPVLAGLFGVAVAPNTEITREIVHSL 120
DB 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFIPVLAGLFGVAVAPNTEITREIVHSL 120
QY 121 GISKPTIVSSSKGLDKVITVQKTAIATITVILDSKVDRGYOSMDNFKKNTPOGFKG 180
DB 121 GISKPTIVSSSKGLDKVITVQKTAIATITVILDSKVDRGYOSMDNFKKNTPOGFKG 180
QY 181 SSFKTEVNRKEQVALIMNSSGSTGLPKGVOLTHENIVRFSHARDPIRYGNOSPGATL 240
DB 181 SSFKTEVNRKEQVALIMNSSGSTGLPKGVOLTHENIVRFSHARDPIRYGNOSPGATL 240
QY 241 TVPFFHGHGFMFTTGLYLCGFRIVMLTKFDEETPLKTLQDYKCSSVILPTLFAILNRS 300
DB 241 TVPFFHGHGFMFTTGLYLCGFRIVMLTKFDEETPLKTLQDYKCSSVILPTLFAILNRS 300
QY 301 ELLDKYDLNIVEIASGAPLSKEIGEAVARRNLCGVNRQVGLTETTSATITTPGDDK 360
DB 301 ELLDKYDLNIVEIASGAPLSKEIGEAVARRNLCGVNRQVGLTETTSATITTPGDDK 360
QY 361 PGASGVVPLFKAKVIDDITDKTKTLGPNRRGEVCVKGPMIMKGVNDPPEATREITDEEGL 420
DB 361 PGASGVVPLFKAKVIDDITDKTKTLGPNRRGEVCVKGPMIMKGVNDPPEATREITDEEGL 420
QY 421 HNGDIGYDEEKNHFIYDRKSLIKYKGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
DB 421 HNGDIGYDEEKNHFIYDRKSLIKYKGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
QY 481 ELPGAUVVLKKGSMTEKEVMDYVASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIKREI 540
DB 481 ELPGAUVVLKKGSMTEKEVMDYVASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIKREI 540
QY 541 LKKPVAKM 548
DB 541 LKKPVAKM 548

RESULT 6
AAR30803
ID AAR30803 standard; Protein: 548 AA.
AC AAR30803;
XX 24-MAY-1993 (first entry)
DE Thermostable luciferase (wild-type).
XX Firefly; luciola cruciata; GENJI; luciola lateralis; HEIKE;
XX mutant; assay.
XX Luciola lateralis.

```

```

XX Key Location/Qualifiers
FH Misc-difference 217
FT /note= "amino acid replaced by hydrophobic amino
FT acid, esp. Ile, Leu or Val, for heat stable
FT mutant"
XX
XX EP524448-A.
XX
XX 27-JAN-1993.
XX
XX 26-JUN-1992: 92EP-0110808.
XX
XX 27-JUN-1991: 91JP-0157117.
XX 29-NOV-1991: 91JP-0317064.
XX 22-MAY-1992: 92JP-0131057.
XX
XX (KIKK ) KIKKOMAN CORP.
XX
XX Elachi N, Naoki K;
XX
XX WPI: 1993-028553/04.
XX N-PSDB; AAQ34745.
XX
XX New DNA sequence of a thermostable luciferase - has the aminoacid
XX at position-217 of luciola cruciata or lateralis luciferase
XX replaced by a hydrophobic aminoacid
XX
XX Claim 10-12: Page 30 + 25-29; 33pp: English.
XX
XX A thermostable luciferase of a firefly, in which an amino acid
XX at the 217 position or an amino acid at the position equiv. to the
XX 217 position of the luciferase of Luciola cruciata (GENJI firefly)
XX of Luciola lateralis (HEIKE firefly) is replaced by a hydrophobic
XX amino acids, esp. Ile, Leu or Val, is claimed. The mutant luciferase
XX is identical in properties to the wild-type luciferase except that
XX it is stable when heated to high temps., e.g. 50 degrees C.
XX The luciferase can be used in assays for e.g. ATP.
XX
XX Sequence 548 AA:
XX
Query Match 99.7%; Score 2814; DB 14; Length 548;
Best Local Similarity 99.6%; Pred. No. 6.2e-235;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 MENNENDENTIVYGPEPPYPIEESGAGOLRKYMRYAKGALFTNALTGVDYTAAYLE 60
XX
XX 1 memmendentivygpeppyleegsagqlrkymdryaklgalfnaltgvdyltaeyle 60
XX
XX 61 KSCCLGKALWGLVVDGRALACSENEEPFIPVLGLFTGVGAPNNEITVLRREIVHSL 120
XX
XX 61 kscclgealkwglvvdgralacseneeefipvlglftgvgapnneitvlelvhsl 120
XX
XX 121 GISKPTIYFSSKGLDKYITVQKTVTAIKTIVILDSVDYKGYSDMNFKKNTPOGFKG 180
XX
XX 121 gisrptiyfsskglkdyitvqktvtaiiktivildsvdykgyssdmnfkkntpogfkg 180
XX
XX 121 gisrptiyfsskglkdyitvqktvtaiiktivildsvdykgyssdmnfkkntpogfkg 180
XX
XX 181 SSFKEVYENRKEQVALMNSGSGGLPKGVOLTHEENTVTRNSARADPIYGNQVSPGAIL 240
XX
XX 181 ssfkevyenrkeqvalmnssgsgglpkgvolttheentvtrnsaradpiygnqvspgail 240
XX
XX 181 ssfkevyenrkeqvalmnssgsgglpkgvolttheentvtrnsaradpiygnqvspgail 240
XX
XX 241 TVVPHHGFGEFTLGLYTCGFRIVMLTKFDEETFLTKIDYKSSVILVPTLFAINRS 300
XX
XX 241 tvvphhgfgeftlgllytcgfrivmltkfdeetfltkidykssviltvptlfaalnrs 300
XX
XX 241 tvvphhgfgeftlgllytcgfrivmltkfdeetfltkidykssviltvptlfaalnrs 300
XX
XX 301 ELDDKYDLSNLEIASGAPLSKETGEAVARRFNLPGVROGYGLTETTSALITTPBGDDK 360
XX
XX 301 elldkydlsnleiasgaplsketgeavarrrfnlpgvrogygltettsalittpebgddk 360
XX
XX 301 elldkydlsnleiasgaplsketgeavarrrfnlpgvrogygltettsalittpebgddk 360
XX
XX 361 PGASGRVPLRKAVYIDDKTKKTLGPNRGEVCVGPMLMGVYDNPATREIIDEEGWL 420
XX
XX 361 pgasgrvplrkavyiddtkkttlgnrgevcvgpmlmgvydnpatreiideegwl 420
XX
XX 361 pgasgrvplrkavyiddtkkttlgnrgevcvgpmlmgvydnpatreiideegwl 420

```

```

XX
XX 421 HTGDIGYVDEKHEFEIVDRILKSLIKYKGYVPAPELESVLLQHPNIFDAGVAPPIAG 480
XX
XX 421 htgdigyvdekhfeivdrilkslikykyvpapeesvllqhp nifdagvappiag 480
XX
XX 481 ELPCAVVVLKKGKSMTEKEVMDYVASQVSNARKRLRGVRVEDEPKGLTGKIDGKAIRET 540
XX
XX 481 elpcavvvllkkgksmtekevmdyvasqvsnakrlrgvrvfedepkgltgkldgkai ret 540
XX
XX 541 LKKPVAKM 548
XX
XX 541 lkkpvakm 548
XX
XX 541 lkkpvakm 548
XX
XX RESULT 7
XX AAR33710
XX ID AAR33710 standard; Protein: 548 AA.
XX
XX AAR33710;
XX
XX 24-MAY-1993 (first entry)
XX
XX Thermostable luciferase (T217L,L,V).
XX
XX Firefly: Luciola cruciata; GENJI; Luciola lateralis; HEIKE;
XX mutant; assay.
XX
XX Luciola lateralis.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 217
XX FT /label= ILE, LEU, VAL
XX
XX EP524448-A.
XX
XX 27-JAN-1993.
XX
XX 26-JUN-1992: 92EP-0110808.
XX
XX 27-JUN-1991: 91JP-0157117.
XX 29-NOV-1991: 91JP-0317064.
XX 22-MAY-1992: 92JP-0131057.
XX
XX (KIKK ) KIKKOMAN CORP.
XX
XX Elachi N, Naoki K;
XX
XX WPI: 1993-028553/04.
XX
XX New DNA sequence of a thermostable luciferase - has the aminoacid
XX at position-217 of luciola cruciata or lateralis luciferase
XX replaced by a hydrophobic aminoacid
XX
XX Claim 10-12: Page 30 + 25-29; 33pp: English.
XX
XX A thermostable luciferase of a firefly, in which an amino acid
XX at the 217 position or an amino acid at the position equiv. to the
XX 217 position of the luciferase of Luciola cruciata (GENJI firefly)
XX of Luciola lateralis (HEIKE firefly) is replaced by a hydrophobic
XX amino acids, esp. Ile, Leu or Val, is claimed. The mutant luciferase
XX is identical in properties to the wild-type luciferase except that
XX it is stable when heated to high temps., e.g. 50 degrees C.
XX The luciferase can be used in assays for e.g. ATP.
XX
XX Sequence 548 AA:
XX
Query Match 99.7%; Score 2814; DB 14; Length 548;
Best Local Similarity 99.6%; Pred. No. 6.2e-235;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 MENNENDENTIVYGPEPPYPIEESGAGOLRKYMRYAKGALFTNALTGVDYTAAYLE 60
XX
XX 1 memmendentivygpeppyleegsagqlrkymdryaklgalfnaltgvdyltaeyle 60
XX
XX 61 KSCCLGKALWGLVVDGRALACSENEEPFIPVLGLFTGVGAPNNEITVLRREIVHSL 120
XX
XX 61 kscclgealkwglvvdgralacseneeefipvlglftgvgapnneitvlelvhsl 120
XX
XX 121 GISKPTIYFSSKGLDKYITVQKTVTAIKTIVILDSVDYKGYSDMNFKKNTPOGFKG 180
XX
XX 121 gisrptiyfsskglkdyitvqktvtaiiktivildsvdykgyssdmnfkkntpogfkg 180
XX
XX 121 gisrptiyfsskglkdyitvqktvtaiiktivildsvdykgyssdmnfkkntpogfkg 180
XX
XX 181 SSFKEVYENRKEQVALMNSGSGGLPKGVOLTHEENTVTRNSARADPIYGNQVSPGAIL 240
XX
XX 181 ssfkevyenrkeqvalmnssgsgglpkgvolttheentvtrnsaradpiygnqvspgail 240
XX
XX 181 ssfkevyenrkeqvalmnssgsgglpkgvolttheentvtrnsaradpiygnqvspgail 240
XX
XX 241 TVVPHHGFGEFTLGLYTCGFRIVMLTKFDEETFLTKIDYKSSVILVPTLFAINRS 300
XX
XX 241 tvvphhgfgeftlgllytcgfrivmltkfdeetfltkidykssviltvptlfaalnrs 300
XX
XX 241 tvvphhgfgeftlgllytcgfrivmltkfdeetfltkidykssviltvptlfaalnrs 300
XX
XX 301 ELDDKYDLSNLEIASGAPLSKETGEAVARRFNLPGVROGYGLTETTSALITTPBGDDK 360
XX
XX 301 elldkydlsnleiasgaplsketgeavarrrfnlpgvrogygltettsalittpebgddk 360
XX
XX 301 elldkydlsnleiasgaplsketgeavarrrfnlpgvrogygltettsalittpebgddk 360
XX
XX 361 PGASGRVPLRKAVYIDDKTKKTLGPNRGEVCVGPMLMGVYDNPATREIIDEEGWL 420
XX
XX 361 pgasgrvplrkavyiddtkkttlgnrgevcvgpmlmgvydnpatreiideegwl 420
XX
XX 361 pgasgrvplrkavyiddtkkttlgnrgevcvgpmlmgvydnpatreiideegwl 420

```

```

Db      1 memmendeniyygpepfyieegsagaqlrkymdryakigatafinaltgvdtyaeyle 60
QY      61 KSCCLGEALKNYGLVVDGRIALCSNCEEFPIVLAGLFVGAVAPRNEYTLRELVHSL 120
        |||
Db      61 kscclgealknyglvvdgrialcsenceeffipvlaglfvgavaprneyltrelvhsl 120
QY      121 GISKPTIVSSSKKGLDVKVTVTQKTVTAIKTVIILDSKVDYRGYSMDNEIKKNTPOGFKG 180
        |||
Db      121 giskptivsskkgldkvltvgktvtaiktviildskvdrygysmdnflkntpgfkf 180
QY      181 SSFKTEVNRKEQVALIMNSSGSTGLPKGVOLTHENTVTRFSHARPIYGNQVSPGTAIL 240
        |||
Db      181 ssfktevnrkeqvalimnssgstglpkvgqlthenvtrfshardpiygnvsgptail 240
QY      241 TVPPEHHGFMFTLGYITGCFRIYMLTKPDEETFLKTLQDYKCSSVILVPTLFAILNRS 300
        |||
Db      241 tvpphhgfmftlgyitlgytcfriymltkpdeetflktdykcssvllvptlfaillrs 300
QY      301 ELDPKVDLSNLVEIASGAPLSKEIGENAVARFNLPGVROGVTGLETTSATITPBGDDK 360
        |||
Db      301 elldkvdlsnlveiasgapslskeigeavarfnlpgvrogtglettsaititpbgddk 360
QY      361 PGASGVVPLFKAKVIDIDTKKTLGNRRGEVCVKGPMIMKGYVNDPEATREIDEEGWL 420
        |||
Db      361 pgasgkvvplfkakvididtkktlgnrrgevcvkgpmlmkgyvndpeatreideegwl 420
QY      421 HTGDIGYDEEKHFIVTRKSLIKYGYVPALESVLLQHPNIFDAGVAGVPPDIAG 480
        |||
Db      421 htgdigydeekhfivtrkslikygyvpaelsevlqhpnifdagvagvppdiag 480
QY      481 ELPGAVVLLKKGSKMTEKEVMDYVASOVSNAKRLRGVRFVDEVPKGLTGKIDGAKIREI 540
        |||
Db      481 elpgavvllkkgksmtekevmdyvasovsnakrlrgvrfvdevpkgltgkldgkairer 540
QY      541 LKKPVAKM 548
        |||
Db      541 lkkpvakm 548
        |||

RESULT 8
AAW12395
ID      AAW12395 standard; Protein: 636 AA.
XX
AC      AAW12395;
XX
DT      20-MAY-1997 (first entry)
XX
DE      Firefly luciferase conjugated with 87 aa biotin peptide at 3' end.
XX
KW      Fusion protein; firefly; luciola lateralis; luciferase; biotinylation;
        wild type; E.coli; bioluminescence assay; quantification; ligand;
        receptor.
XX
OS      Synthetic.
XX
FH      Key
FT      Region
        1..547
        /note="Firefly luciferase portion of fusion protein"
FT      Region
        548..549
        /note="encoded by linker sequence used to construct
        fusion gene"
FT      Region
        549..636
        /note="biotin peptide portion of fusion protein"
XX
PN      JP08308578-A.
XX
PD      26-NOV-1996.
XX
PF      24-APR-1995; 95JP-0098857.
XX
PR      14-MAR-1995; 95JP-0054625.
XX
PR      27-JUL-1994; 94JP-0193798.
XX

```

```

PA      (KIKK ) KIKKOMAN CORP.
XX
XX      WPI: 1997-059697/06.
DR      N-PSDB; AAT63269.
XX
PT      Fusion protein comprising firefly luciferase and biotinylated
        peptide - useful in a bio-luminescent analytical method for
        quantifying ligands
XX
XX      Claim 2; Page 12-13; 13pp; Japanese.
XX
XX      This is the amino acid sequence of a novel fusion protein which comprises
        the firefly (Luciola lateralis) luciferase protein and an 87 amino acid
        biotinylated peptide designated the E.coli biotin carboxy-carrier protein
        (BCCP-87). The fusion gene was generated by firstly amplifying the
        sequence encoding the E.coli BCCP-87 from the E.coli genome and inserting
        the resultant sequence into the 3' end of the firefly luciferase gene in
        plasmid pHLf230. The biotin peptide sequence replaces the C-terminal Met
        residue of the wild type luciferase sequence. The plasmid was then
        introduced into E.coli JM101 for production of the fusion protein. The
        luciferase ligands which may modulate binding of luciferase to its
        receptor.
CC
CC      Sequence 636 AA;
SQ

```

```

Query Match          99.6%; Score 2812; DB 18; Length 636;
Best Local Similarity 99.6%; Pred. No. 1,1e-234;
Matches 545; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 MEMMENDENIYYGPEPFYIEEGSAGAO LRKYMDRYAKIGATAFNALTGVDYTYAEYLE 60
        |||
Db      1 memmendeniyygpepfyieegsagaqlrkymdryakigatafinaltgvdtyaeyle 60
QY      61 KSCCLGEALKNYGLVVDGRIALCSNCEEFPIVLAGLFVGAVAPRNEYTLRELVHSL 120
        |||
Db      61 kscclgealknyglvvdgrialcsenceeffipvlaglfvgavaprneyltrelvhsl 120
QY      121 GISKPTIVSSSKKGLDVKVTVTQKTVTAIKTVIILDSKVDYRGYSMDNEIKKNTPOGFKG 180
        |||
Db      121 giskptivsskkgldkvltvgktvtaiktviildskvdrygysmdnflkntpgfkf 180
QY      181 SSFKTEVNRKEQVALIMNSSGSTGLPKGVOLTHENTVTRFSHARPIYGNQVSPGTAIL 240
        |||
Db      181 ssfktevnrkeqvalimnssgstglpkvgqlthenvtrfshardpiygnvsgptail 240
QY      241 TVPPEHHGFMFTLGYITGCFRIYMLTKPDEETFLKTLQDYKCSSVILVPTLFAILNRS 300
        |||
Db      241 tvpphhgfmftlgyitlgytcfriymltkpdeetflktdykcssvllvptlfaillrs 300
QY      301 ELDPKVDLSNLVEIASGAPLSKEIGENAVARFNLPGVROGVTGLETTSATITPBGDDK 360
        |||
Db      301 elldkvdlsnlveiasgapslskeigeavarfnlpgvrogtglettsaititpbgddk 360
QY      361 PGASGVVPLFKAKVIDIDTKKTLGNRRGEVCVKGPMIMKGYVNDPEATREIDEEGWL 420
        |||
Db      361 pgasgkvvplfkakvididtkktlgnrrgevcvkgpmlmkgyvndpeatreideegwl 420
QY      421 HTGDIGYDEEKHFIVTRKSLIKYGYVPALESVLLQHPNIFDAGVAGVPPDIAG 480
        |||
Db      421 htgdigydeekhfivtrkslikygyvpaelsevlqhpnifdagvagvppdiag 480
QY      481 ELPGAVVLLKKGSKMTEKEVMDYVASOVSNAKRLRGVRFVDEVPKGLTGKIDGAKIREI 540
        |||
Db      481 elpgavvllkkgksmtekevmdyvasovsnakrlrgvrfvdevpkgltgkldgkairer 540
QY      541 LKKPVAK 547
        |||
Db      541 lkkpvak 547
        |||

RESULT 9

```



```

Db      61 ksccljealknyglvvdgrtalcsenceeffirvlaglfigvavaptnelylrelvhs1 120
QY      121 GISKPTIVFSSKKGLDKVITVQKTVTAIKRTIVLDSKVDRGQSDMNFIKKMTPOGFKG 180
XX      |||||||
XX      122 glskprlvtsskkgldkvlvtqktvtaikrtivldskvdrygysmdnflkntlpqfkg 180
QY      181 SSKRTVEVNRKEQVALIMNSSGSTGLPKGVQLTHENIVTRESHARDPIYGNQVSPGTAIL 240
XX      |||||||
XX      181 sskrtvevnrkeqvalimnssgstglpkgyqlthenavtrfshardpiygnvsgptail 240
QY      241 IVVPEFHGFGMFTLLGYLTCGFRIVMLTKRDEPTFLKTIQDYKSSVLIIVPTLAILNRS 300
XX      |||||||
XX      241 tvpfhngfgmftllgyllcgfrivmltkfdeeflklqdykssvliivptlailnrs 300
QY      301 ELIDKYDLSNLVEIASGAPLSKEIGEAARFNLPGVRQGYGLTETTSAILIITPEGDDK 360
XX      |||||||
XX      301 elldkydlslnlveiasgapslskeigeavarrrfnlpgvrqgylltettssailitpegddk 360
QY      361 PGASGVVPLFKAKVITDITKTKLGNRRGVEVCVKPMIMKGYVNDPEATREIIDEGBWL 420
XX      |||||||
XX      361 pgasgvvplfkakvilditkktlgnrrgevcvkgpmlmkgyvndpeatreiideegw1 420
QY      421 HFGDIDGYDEEKNHFTVDRLKSLIKYKGVOVPAELSEVILQHPNIFDAGVAVPPPIAG 480
XX      |||||||
XX      421 hfgdidgydeeknhftvdrlkslikykgvoppaelsevilqhpnlfdagvavpppiag 480
QY      481 ELPGAVVVLKKGKSMTEKEVMDYVASQVSNARKLRGVRFEVDEVPKGLTGKIDGKAIRREI 540
XX      |||||||
XX      481 elpgavvvllkkgksmtekevmdyvasqvsnakrlrgvrfdevdpkgltgkldgkairrel 540
QY      541 LKKPVAKM 548
XX      |||||||
XX      541 lkkpvakm 548
Db

```

RESULT 11
AAM12394
ID AAM12394 standard; Protein: 568 AA.
AC AAM12394;
XX
XX 20-MAY-1997 (first entry)
XX
XX Firefly luciferase conjugated with 23 aa biotin peptide at N-terminus.
DE
XX Fusions protein; firefly; luciola lateralis; luciferase; biotinylation;
KW wild type; E.coli; bioluminescence assay; quantification; ligand;
KW receptor.
XX
XX Synthetic.
OS
FH
FH Key Location/Qualifiers
FT 1..24
FT /note= "biotin peptide"
FT 25..568
FT Region /note= "firefly luciferase"
FT
XX JP08308578-A.
XX
XX 26-NOV-1996.
XX
XX 24-APR-1995; 95JP-0098857.
XX
XX 14-MAR-1995; 95JP-0054625.
XX 27-JUL-1994; 94JP-0193798.
XX
XX (KIRK) KIRKMAN CORP.
XX
XX WPI: 1997-059697/06.
XX N-PSDB: AAT63268.
XX
XX Fusion protein comprising firefly luciferase and biotinylated
PT peptide - useful in a bio-luminescent analytical method for

```

PT      quantifying ligands
XX
XX Claim 2; Page 10-11; 13pp; Japanese.
XX
XX This is the amino acid sequence of a novel fusion protein comprising
CC the firefly (luciola lateralis) luciferase protein and a 23 amino acid
CC biotinylated peptide designated biotin peptide #84. The gene was
CC generated by firstly annealing the sequence encoding peptide #84
CC (AAT63270) and its complementary sequence, then inserting the resultant
CC double stranded sequence into the 5' end of the luciferase gene in
CC plasmid pHL108. The biotin peptide sequence replaces the first 4 amino
CC acids of the wild type luciferase sequence. The plasmid was then
CC introduced into E.coli JM101 for production of the fusion protein. The
CC novel protein can be used in bioluminescence assays to quantify
CC luciferase ligands which may modulate binding of luciferase to its
CC receptor.
SQ      Sequence 568 AA:

```

Query Match 99.1%; Score 2799; DB 18; Length 568;
Best Local Similarity 99.3%; Pred. No. 1.3e-233;
Matches 542; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

QY      3 NMENDENIYGPPEPPIYIEGSAQOLRRKYMRYAKLGAIAFTNALTGVDYTAEYLEKS 62
XX      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
XX      23 slendenivgyppfpfyiegsaqqlrkymdryaklgatalfnaltgdydyaeyleks 82
QY      63 CCLGEALKNYGLVNDORIALCSENCEEFIPVLAGFIVGAVAPNTEITLRELVHSLGI 122
XX      |||||||
XX      63 cclgealknvglvndorialsenceeffipvlagfigvavapnteitlrelvhs1g1 142
QY      83 CCLGEALKNYGLVNDGRIALCSENCEEFFIPVLAGFIVGAVAPNTEITLRELVHSLGI 142
XX      |||||||
XX      83 cclgealknvglvndgrialsenceeffipvlagfigvavapnteitlrelvhs1g1 142
QY      123 SKPTIVFSSKKGLDKVITVQKTVTAIKRTIVLDSKVDRGQSDMNFIKKMTPOGFKGSS 182
XX      |||||||
XX      143 skptivfsskkgldkvlvtqktvtaikrtivldskvdrygysmdnflkntlpqfkgss 202
QY      183 FKTVEVNRKEQVALIMNSSGSTGLPKGVQLTHENIVTRESHARDPIYGNQVSPGTAILTV 242
XX      |||||||
XX      203 fktvevnrkeqvalimnssgstglpkgyqlthenavtrfshardpiygnvsgptailtv 262
QY      243 VPFHNGFGMFTLLGYLTCGFRIVMLTKRDEPTFLKTIQDYKSSVLIIVPTLAILNRSLL 302
XX      |||||||
XX      263 vpfhngfgmftllgyllcgfrivmltkfdeeflklqdykssvliivptlailnrsel 322
QY      303 LDKYDLSNLVEIASGAPLSKEIGEAARFNLPGVRQGYGLTETTSAILIITPEGDDKPG 362
XX      |||||||
XX      323 ldkydlslnlveiasgapslskeigeavarrrfnlpgvrqgylltettssailitpegddkpg 382
QY      363 ASGKVYVPLFKAKVITDITKTKLGNRRGVEVCVKPMIMKGYVNDPEATREIIDEGBWLHT 422
XX      |||||||
XX      383 asgkvvplfkakvilditkktlgnrrgevcvkgpmlmkgyvndpeatreiideegw1ht 442
QY      423 GDIGYDEEKNHFTVDRLKSLIKYKGVOVPAELSEVILQHPNIFDAGVAVPPPIAGRL 482
XX      |||||||
XX      443 gdigydeeknhftvdrlkslikykgvoppaelsevilqhpnlfdagvavpppiagrl 502
QY      483 PGAVVVLKKGKSMTEKEVMDYVASQVSNARKLRGVRFEVDEVPKGLTGKIDGKAIRREILK 542
XX      |||||||
XX      503 pgavvvllkkgksmtekevmdyvasqvsnakrlrgvrfdevdpkgltgkldgkairrelk 562
QY      543 KPVAKM 548
XX      |||||||
XX      543 kpvakm 568
Db      563 KPVAKM 568

```

RESULT 12
AAM53883
ID AAM53883 standard; Protein: 568 AA.
AC AAM53883;
XX
XX 16-JUL-1998 (first entry)
XX

XX	Antibody-Firefly Luciferase fusion protein.
XX	Firefly Luciferase; antibody-Luciferase fusion protein.
KM	
XX	Synthetic.
OS	Luciola cruciata.
XX	
XX	JP09187281-A.
PN	
XX	
PD	22-JUL-1997.
XX	
PF	09-JAN-1996; 96JP-0001812.
XX	
XX	09-JAN-1996; 96JP-0001812.
PR	
XX	
PA	(KIRK) KIRKMAN CORP.
XX	
DR	WPI; 1998-275089/25.
DR	N-PSDB; AAV23595.
XX	
PT	Antibody-firefly Luciferase fused protein - and related products
PT	i.e. firefly luciferase fused gene, recombinant DNA and its
PT	preparation
XX	
XX	Disclosure: Page 11-13; 17pp; Japanese.
XX	
CC	This sequence is a fusion protein of the invention. The protein is
CC	a antibody-firefly luciferase fusion protein, in which an antibody part
CC	consisting of a peptide having antibody activity is combined with an
CC	enzyme part consisting of firefly luciferase.
XX	
SQ	Sequence 568 AA.

Query Match	99.1%	Score 2799	DB 19	Length 568
Similarity		Pred. No. 1.3e-233		
Best Local	99.3%			
Matches 542; Conservative	4	Mismatches 0	Indels 0	Gaps 0

[illegible]

QY	543	KPVAKM	548
Db	563	KPVAKM	568

RESULT	13
AAW04208	
ID	AAW04208 standard; Protein; 673 AA

DT	12-NOV-1996	(first entry)
XX		
DE		Streptavidin/Luciferase fusion protein.
XX		
KM		Streptavidin: luciferase: fusion protein;
KM		Streptomyces avidinii: lucifera lateralis; firefly
KM		recombinant production; industry.

FT	Key	Location/Qualifiers
FT	Protein	1..128
FT		/label= streptavidin
FT	Protein	130..673
FT		/label= luciferase

PN JP07289264-A.

PD 07-NOV-1995

PF 27-APR-1994; 94JP-0090275.

PR 27-APR-1994; 94JP-0090275.

PA (KIKK) KIKKOMAN CORP

DR WPI; 1996-015269/02.
DR N-PSDB; AAT33850.

XX New mutant streptavidin (SA) gene - fused to a firefly luciferase
 PT gene, for the recombinant prepn. of a SA-Fl fused protein
 XX
 PS Example 1; Pages 6-7; 12pp; Japanese.

The present sequence is a streptavidin/luciferase fusion protein, comprising the Streptococcus avidin streptavidin gene and the Lucifera lateralis (firefly) luciferase gene. The fusion protein can be prepri by inserting the recombinant DNA encoding it into a Escherichia species microbe, culturing the transformed microbe in a medium and collecting the fusion protein from the CC culture. The fusion protein can be used in industry.

Sequence	673 AA
5Q	

Query Match	99.1%	Score 2799;	DB 17;	Length 673;
Best Local Similarity	99.3%;	Pred. No. 1,7e-233		
Matches 542; Conservative	4;	Mismatches 0;	Indels 0;	Gaps 0

QY	3	NMENDENIVGPEPFPIVIEGSGAGAOBKRYMDRYAKLGAIATNANLGVDDTVEYIEKS	62
:	:		:
Db	128	slendenilvygpepfypieegsagaqlrkymdryaklgafatnaltpvdylyeayleks	187
QY	63	CCLEALKNTGVLVDRGIALCSENGCEEPFLVLAGLFIQVGAAPNEIYTLTRELVHSLGI	122
Db	188	cclealknqylvvdgrfialcseenceffltpvaqlflisvgyapnelylrlrelvhsqyl	247
QY	123	SKPTIVSSKKGLDKVITVQKTYTAIKTIVILIDSKVDYRGVOSMDNFIKNTPOGFGSS	182
Db	248	skptivfsskkqldkvltvqkvtatiktilyldskvdgygysmdnfiikntbpgfkfys	307

```

OY 183 FRTVEVNRKEVOALIMNSSGSTGLPKGVOLTHEENITVRFESHARDPIYGNVSPGTAILTV 242
DB 308 fktvevnrkegvalimnssgstglpkgyvqlthenivtrfshardpiygnvspgtailtv 367
OY 243 VPFHHGFMFTTGLYLTGCFRIYMLTKFDEETFLKLDQYKSSVILVPTLFAILNRSSEL 302
DB 368 vpfhngfgmfttlyltcgfrilvmltkideetflkltdqykssviltplfaalnrsel 427
OY 303 LDKYDLSNLVEIASGAPLSKEIGEAVARRNLPGRVGYLTETTSAILITPBGDDKPG 362
DB 428 ldkydlslnlveliasgapslskeigearrlnlpgrvgylltetsailitpbgddkpg 487
OY 363 ASGKVPPLFKAKVIDLDTKRTTLGPNRGEVCVKGMKMGVYDNPDEARETIDEEGMLHT 422
DB 488 asgkvvplfkakvidltdkrttlgpnrrgevcvkgmllmkgyvndpearetreideegwllht 547
OY 423 GDIGYDEEKHFPIVDRISLTKIKYGVPPAELESVLLQHPNIFDAGVAVPPIAGEL 482
DB 548 gdigydeekhfivdrilsklkygvpaelesvllqhpnlfdagvagvpdpiagel 607
OY 483 PGAVVVLKKGKSMTEKEVMYVASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIREILK 542
DB 608 pgavvvlekgsmtkekemdyvasqvsnakrlrgvrfvdevpklgkldgkaidreilk 667
OY 543 KPVAKM 548
DB 668 kpvakm 673

RESULT 14
ID AAM04209
AC AAM04209;
XX
XX 12-NOV-1996 (first entry)
DE Mutant streptavidin/luciferase fusion protein.
XX
XX Mutant; streptavidin; luciferase; fusion protein;
KM Streptomyces avidinii; Luciola lateralis; firefly;
KW recombinant production; industry.
XX
XX Synthetic.
XX
XX Key
FH 1.140 Location/Qualifiers
FT Protein /label= mutant streptavidin
FT Protein /label= luciferase
FT Peptide 1.17
FT /label= mutant N-terminal peptide

JP07289264-A.
XX
XX 07-NOV-1995.
XX
XX 27-APR-1994; 94JP-0090275.
XX
XX 27-APR-1994; 94JP-0090275.
XX
XX (KIRK ) KIRKMAN CORP.
XX
XX WPI; 1996-015269/02.
XX DR N-PSDB; AAT33851.
XX
XX New mutant streptavidin (SA) gene - fused to a firefly luciferase
XX PT gene, for the recombinant prepn. of a SA-FL fused protein
XX
XX Claim 4; Pages 8-9; 12pp; Japanese.
XX
XX The present sequence is a mutant streptavidin/luciferase fusion
CC protein, comprising a mutant Streptomyces avidinii streptavidin and
CT

```

```

CC the Luciola lateralis (firefly) luciferase. The fusion protein can
CC be prep'd. by inserting the recombinant DNA encoding it into a
CC Escherichia species microbe, culturing the transformed microbe in a
CC medium and collecting the fusion protein from the culture. The
CC fusion protein can be used in industry.
XX
XX Sequence 685 AA:

Query Match 99.1%; Score 2799; DB 17; Length 685;
Best Local Similarity 99.3%; Pred. No. 1.7e-233;
Matches 542; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 3 NMENDENIYGPPEPFPIEGSAGAO LRKYMDRYAKGAIATFNALTGDYTYARELEKS 62
DB 140 slendenilvygpepfpiiegsagaglrlymtryaklgatlnaltgydyaeyleks 199
OY 63 CCLGELAKNYGLVYDRIALCSENCEEFIPVLAGLFIGVGAAPNIEIYTLRELVHSLGI 122
DB 200 ccclgealnknyglvdydrialcsenceeffipvlaglfigvgapnelytlrelvhsigli 259
OY 123 SKPTIVFSSKRGIDKVIYQKTVTAIKRTIIVLDSKVDYRGYQSMNFIKKNPQGRKSS 182
DB 260 skptivfsskrgldkvtvqkvtvtaikrtivlidskvdrygyqsmnflkntpqgrkss 319
OY 183 FRTVEVNRKEVOALIMNSSGSTGLPKGVOLTHEENITVRFESHARDPIYGNVSPGTAILTV 242
DB 320 fktvevnrkegvalimnssgstglpkgyvqlthenivtrfshardpiygnvspgtailtv 379
OY 243 VPFHHGFMFTTGLYLTGCFRIYMLTKFDEETFLKLDQYKSSVILVPTLFAILNRSSEL 302
DB 380 vpfhngfgmfttlyltcgfrilvmltkideetflkltdqykssviltplfaalnrsel 439
OY 303 LDKYDLSNLVEIASGAPLSKEIGEAVARRNLPGRVGYLTETTSAILITPBGDDKPG 362
DB 440 ldkydlslnlveliasgapslskeigearrlnlpgrvgylltetsailitpbgddkpg 499
OY 363 ASGKVPPLFKAKVIDLDTKRTTLGPNRGEVCVKGMKMGVYDNPDEARETIDEEGMLHT 422
DB 500 asgkvvplfkakvidltdkrttlgpnrrgevcvkgmllmkgyvndpearetreideegwllht 559
OY 423 GDIGYDEEKHFPIVDRISLTKIKYGVPPAELESVLLQHPNIFDAGVAVPPIAGEL 482
DB 560 gdigydeekhfivdrilsklkygvpaelesvllqhpnlfdagvagvpdpiagel 619
OY 483 PGAVVVLKKGKSMTEKEVMYVASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIREILK 542
DB 620 pgavvvlekgsmtkekemdyvasqvsnakrlrgvrfvdevpklgkldgkaidreilk 679
OY 543 KPVAKM 548
DB 680 kpvakm 685

RESULT 15
ID AAM53882
AC AAM53882;
XX
XX 16-JUL-1998 (first entry)
DE Antibody-Firefly luciferase fusion protein.
XX
XX Firefly luciferase; antibody-luciferase fusion protein.
XX
XX Synthetic.
XX OS Luciola cruciata.
XX
XX Key
FH Location/Qualifiers
FT Misc-difference 224
FT /note= "encoded by CAG"
FT Misc-difference 247

```


QY 241 TVPFFHGFMTTGLYLTGFRIVMLTKPDEETFLKTLDDYKCSSVILVPTLAILNRS 300
 |||||
 Db 241 TVPFFHGFMTTGLYLTGFRIVMLTKPDEETFLKTLDDYKCSSVILVPTLAILNRS 300
 QY 301 ELLDKYDLSNVLKASGAPLSKEIGAVARFNLPGVROGYGLTETTSAILITPEGDGK 360
 |||||
 Db 301 ELLDKYDLSNVLKASGAPLSKEIGAVARFNLPGVROGYGLTETTSAILITPEGDGK 360
 QY 361 PGASGVVPLFKAKVIDDITKTLGPNRRGEVCVKGPMIMKGVDNPEATREIIDEEGWL 420
 |||||
 Db 361 PGASGVVPLFKAKVIDDITKTLGPNRRGEVCVKGPMIMKGVDNPEATREIIDEEGWL 420
 QY 421 HTGDIGYDEEKHFIVDRLSLTKYKGYVPPAELESVLLQHPNIFDAGVAGVPPPIAG 480
 |||||
 Db 421 HTGDIGYDEEKHFIVDRLSLTKYKGYVPPAELESVLLQHPNIFDAGVAGVPPPIAG 480
 QY 481 ELPGAVVLLKKGSMTEKEVMDYVASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
 |||||
 Db 481 ELPGAVVLLKKGSMTEKEVMDYVASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
 QY 541 LKRPVAKM 548
 |||||
 Db 541 LKRPVAKM 548

RESULT 2
 J50181
 Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - Genj1 firefly

N:Alternate names: firefly luciferase
 C:Species: Luciola cruciata (Genj1 firefly)
 C>Date: 07-Jun-1990 #sequence_revision 20-Aug-1994 #text_change 26-May-2000
 C:Accession: J50181
 R:Masuda, T.; Tatsumi, H.; Nakano, E.
 Gene 77, 265-270, 1989
 A:Title: Cloning and sequence analysis of cDNA for luciferase of a Japanese firefly, Luc
 A:Reference number: J50181; MUID:89326143
 A:Accession: J50181
 A:Molecule type: mRNA
 A:Residues: 1-548 <MAS>
 A:Cross-references: GB:M6194; NID:9159050; PIDN:AAB29135.1; PID:9159051
 C:Comment: This protein catalyzes the oxidation of luciferin in the presence of ATP, oxy
 species from green to yellow.
 C:Superfamily: 4-comumate--CoA ligase; acetate--CoA ligase homology
 C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase
 F:73-537/Domain: acetate--CoA ligase homology <ACL>

Query Match 95.1%; Score 2684; DB 1; Length 548;
 Best Local Similarity 93.4%; Pred. No. 2,1e-183;

Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

QY 1 MENMENDENIVYGPPEPYPIEESAGAOIRKYMRYAKAGIAFTNALGVDTYAEYLE 60
 |||||
 Db 1 MENMENDENIVYGPPEPYPIEESAGAOIRKYMRYAKAGIAFTNALGVDTYAEYLE 60
 QY 61 KSCCGEALKNGLVVDGRIALCSENCEFFIPVLAGLFGVAVPNEIYTLRELHVS 120
 |||||
 Db 61 KSCCGEALKNGLVVDGRIALCSENCEFFIPVLAGLFGVAVPNEIYTLRELHVS 120
 QY 121 GISKPTIYSSKKGLDKYITVQKTAIYIILDSKDYRGYQSDNFIKNTPQGFPG 180
 |||||
 Db 121 GISKPTIYSSKKGLDKYITVQKTAIYIILDSKDYRGYQSDNFIKNTPQGFPG 180
 QY 181 SSFKEVVEVNRKEQVALINSSSGSTGLPKGYOLTHENLVTRFSHARDPIYGNQVSGTA 240
 |||||
 Db 181 SSFKEVVEVNRKEQVALINSSSGSTGLPKGYOLTHENLVTRFSHARDPIYGNQVSGTA 240
 QY 241 TVPFFHGFMTTGLYLTGFRIVMLTKPDEETFLKTLDDYKCSSVILVPTLAILNRS 300
 |||||
 Db 241 TVPFFHGFMTTGLYLTGFRIVMLTKPDEETFLKTLDDYKCSSVILVPTLAILNRS 300
 QY 301 ELLDKYDLSNVLKASGAPLSKEIGAVARFNLPGVROGYGLTETTSAILITPEGDGK 360
 |||||

|||
 Db 301 ELLDKYDLSNVLKASGAPLSKEIGAVARFNLPGVROGYGLTETTSAILITPEGDGK 360
 QY 361 PGASGVVPLFKAKVIDDITKTLGPNRRGEVCVKGPMIMKGVDNPEATREIIDEEGWL 420
 |||||
 Db 361 PGASGVVPLFKAKVIDDITKTLGPNRRGEVCVKGPMIMKGVDNPEATREIIDEEGWL 420
 QY 421 HTGDIGYDEEKHFIVDRLSLTKYKGYVPPAELESVLLQHPNIFDAGVAGVPPPIAG 480
 |||||
 Db 421 HTGDIGYDEEKHFIVDRLSLTKYKGYVPPAELESVLLQHPNIFDAGVAGVPPPIAG 480
 QY 481 ELPGAVVLLKKGSMTEKEVMDYVASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
 |||||
 Db 481 ELPGAVVLLKKGSMTEKEVMDYVASQVSNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
 QY 541 LKRPVAKM 548
 |||||
 Db 541 LKRPVAKM 548

RESULT 3
 S33788
 Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) [validated] - sou

N:Alternate names: firefly luciferase
 C:Species: Luciola mingrellica (southern Russian firefly)
 C>Date: 02-Dec-1993 #sequence_revision 03-Nov-1995 #text_change 02-Jun-2000
 C:Accession: S33788
 R:Devine, J.H.; Kutuzova, G.D.; Green, V.A.; Ugarova, N.N.; Baldwin, T.O.
 Biochim. Biophys. Acta 1173, 121-132, 1993
 A:Title: Luciferase from the East European firefly Luciola mingrellica: cloning and nu
 A:Reference number: S33788; MUID:93277944
 A:Accession: S33788
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-548 <DEV>
 A:Cross-references: GB:S61961; NID:9409316; PIDN:AAB26932.1; PID:9409317
 C:Superfamily: 4-comumate--CoA ligase; acetate--CoA ligase homology
 C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase
 F:72-537/Domain: acetate--CoA ligase homology <ACL>

Query Match 82.7%; Score 2335.5; DB 2; Length 548;
 Best Local Similarity 81.7%; Pred. No. 1,4e-158;

Matches 446; Conservative 47; Mismatches 52; Indels 1; Gaps 1;

QY 4 MENMENDENIVYGPPEPYPIEESAGAOIRKYMRYAKAGIAFTNALGVDTYAEYLEKSC 63
 |||||
 Db 3 MEKENYVYGPPEPYPIEESAGAOIRKYMRYAKAGIAFTNALGVDTYAEYLEKSC 62
 QY 64 CLGEALKNGLVVDGRIALCSENCEFFIPVLAGLFGVAVPNEIYTLRELHVS 123
 |||||
 Db 63 RLAEMKNFGMKPEDEHIALCSENCEFFIPVLAGLFGVAVPNEIYTLRELHVS 122
 QY 124 KPTIYSSKKGLDKYITVQKTAIYIILDSKDYRGYQSDNFIKNTPQGFPGSS 183
 |||||
 Db 123 QPTIYSSKKGLDKYITVQKTAIYIILDSKDYRGYQSDNFIKNTPQGFPGSS 182
 QY 184 KTEV-NRKEQVALINSSSGSTGLPKGYOLTHENLVTRFSHARDPIYGNQVSGTA 242
 |||||
 Db 183 VPIDVKKRKHVALLINSSSGSTGLPKGYOLTHENLVTRFSHARDPIYGNQVSGTA 242
 QY 243 VPFHGFMTTGLYLTGFRIVMLTKPDEETFLKTLDDYKCSSVILVPTLAILNRS 302
 |||||
 Db 243 VPFHGFMTTGLYLTGFRIVMLTKPDEETFLKTLDDYKCSSVILVPTLAILNRS 302
 QY 303 LDKYDLSNVLKASGAPLSKEIGAVARFNLPGVROGYGLTETTSAILITPEGDGK 362
 |||||
 Db 303 LDKYDLSNVLKASGAPLSKEIGAVARFNLPGVROGYGLTETTSAILITPEGDGK 362
 QY 363 ASGKVPLFKAKVIDDITKTLGPNRRGEVCVKGPMIMKGVDNPEATREIIDEEGWL 422
 |||||
 Db 363 ASGKVPLFKAKVIDDITKTLGPNRRGEVCVKGPMIMKGVDNPEATREIIDEEGWL 422

S29354

Photinus-luciferin 4-monoxygenase (ATP-hydrolyzing) (EC 1.13.12.7) YE [validated] - Jun

A:Title: Complementary DNA coding click beetle luciferases can elicit bioluminescence

A:Reference number: S29352; MUID:89242142

A:Accession: S29353

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-543 <MOO>

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; monoxygenase; oxidoreductase

F:70-532/Domain: acetate--CoA ligase homology <ACLD>

A:Reference number: S29352; MUID:89242142

A:Accession: S29354

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-543 <MOO>

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; monoxygenase; oxidoreductase

F:70-532/Domain: acetate--CoA ligase homology <ACLD>

Query Match

48.4%; Score 1365; DB 2; Length 543;

Best local similarity 48.2%; Pred. No. 1.9e-89;

Matches 262; Conservative 109; Mismatches 167; Indels 6; Gaps 3;

4 MENDENIVGPEPPYPIEESAGQALRRKYMDRYAKLGAIAFTNLTGVDTYAELEKSC 63

2 MKREKNVYGEPEPLHEDLKTAGEMLFRALKHSHLPQ-ALVDVFGDESLSYKEFEATC 60

64 CLGALKNYGLVVDGRIALSCENCEEFFIPVLAGLFIGVGAAPNFIYTLRELVHSGIS 123

61 LLAGSLNCGYKMDVYVSTICENKREFPIIAMIYGMIVAPYNESYIPDELCKVMSIS 120

124 KPTIVSSKGLDKVITVQKTAIKTIVILDSKVYRGYQSDNFTKNTPOGFKSSSF 183

121 KPQIVFCTKNLKNLVLSVOSTNFKRIIILDVENIHGCSLNFISRYSDGNI--ANF 178

184 KTYEVNKEQVALIMNSGSGTGLPKGYQLTHENLVTRSHARDPIYGNVSPGAILTVV 243

179 KPLHYDVEQVAALICSSGTTGLPKGYQTHQNCVRLIHLADEAGTQLIPGTVLVV 238

244 PFHGFMEFTLLGYLTCGFRIVMLTKFDEETFLTKDQYSSVILVPTLAILNSEL 303

239 PFHAFGFSINLGYFMGLRIVMLRREFQEAFLKAIDYEVRSIVNPAILLFSKSLV 298

304 DKYDLSNVLIASGAPLSKEIGAVARFNLPGVRQGYGLTETTSALITPEBDDKPGA 363

299 DKYDLSIRELCCGAPLAKEVAIAVKRLNLPGRGFGLESTSANIHSLGDEFKSGS 358

364 SGKVPPLFKAKVIDLDTKRTGLPNRGEVCYKGMIMKGYDNEATREITDEGWLHTG 423

359 LGRVTPMAAKIADRETKALGPQVGEELCKGPMVSGYNNVETAKREALIDDGWLHSG 418

424 DIGYDEKHEFFIVDRLSLKIKYGYVPALESLVLQHPNIFDAGVAGVDPDIAGELP 483

419 DFGYDEDEHHYVDRKELIKYGSQVAPALEEILKNPCIDAVAVGIPDLAEGELP 478

484 GAVVVLKKGSMTEKEMVDYVAVSOVSNAKRLRGVRFVDEVPKGLTGKIDGKAIRELK 543

479 SAFVVKOPGKEITAKYDYLAERVSHTKYLRGCVRFVDSIPRVVTGKITRK---ELLKQ 535

544 PVAK 547

536 LLEK 539

RESULT 7

Photinus-luciferin 4-monoxygenase (ATP-hydrolyzing) (EC 1.13.12.7) YG [validated] - Jun

N:Alternate names: firefly-type luciferase

C:Species: Pyrophorus plagiophthalmus

C:Date: 25-Feb-1994 #sequence-revision 01-Sep-1995 #text-change 02-Jun-2000

C:Accession: S29353

R:Wood, K.V.; Lam, Y.A.; Seliger, H.H.; McElroy, W.D.

Science 244, 700-702, 1989

A:Title: Complementary DNA coding click beetle luciferases can elicit bioluminescence

A:Reference number: S29352; MUID:89242142

A:Accession: S29353

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-543 <MOO>

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; monoxygenase; oxidoreductase

F:70-532/Domain: acetate--CoA ligase homology <ACLD>

Query Match

48.4%; Score 1365; DB 2; Length 543;

Best local similarity 48.9%; Pred. No. 1.9e-89;

Matches 266; Conservative 104; Mismatches 168; Indels 6; Gaps 3;

4 MENDENIVGPEPPYPIEESAGQALRRKYMDRYAKLGAIAFTNLTGVDTYAELEKSC 63

2 MKREKNVYGEPEPLHEDLKTAGEMLFRALKHSHLPQ-ALVDVFGDESLSYKEFEATC 60

64 CLGALKNYGLVVDGRIALSCENCEEFFIPVLAGLFIGVGAAPNFIYTLRELVHSGIS 123

61 LLAGSLNCGYKMDVYVSTICENKREFPIIAMIYGMIVAPYNESYIPDELCKVMSIS 120

124 KPTIVSSKGLDKVITVQKTAIKTIVILDSKVYRGYQSDNFTKNTPOGFKSSSF 183

121 KPQIVFCTKNLKNLVLSVOSTNFKRIIILDVENIHGCSLNFISRYSDGNI--ANF 178

184 KTYEVNKEQVALIMNSGSGTGLPKGYQLTHENLVTRSHARDPIYGNVSPGAILTVV 243

179 KPLHYDVEQVAALICSSGTTGLPKGYQTHQNCVRLIHLADEAGTQLIPGTVLVV 238

244 PFHGFMEFTLLGYLTCGFRIVMLTKFDEETFLTKDQYSSVILVPTLAILNSEL 303

239 PFHAFGFSINLGYFMGLRIVMLRREFQEAFLKAIDYEVRSIVNPAILLFSKSLV 298

304 DKYDLSNVLIASGAPLSKEIGAVARFNLPGVRQGYGLTETTSALITPEBDDKPGA 363

299 DKYDLSIRELCCGAPLAKEVAIAVKRLNLPGRGFGLESTSANIHSLGDEFKSGS 358

364 SGKVPPLFKAKVIDLDTKRTGLPNRGEVCYKGMIMKGYDNEATREITDEGWLHTG 423

359 LGRVTPMAAKIADRETKALGPQVGEELCKGPMVSGYNNVETAKREALIDDGWLHSG 418

424 DIGYDEKHEFFIVDRLSLKIKYGYVPALESLVLQHPNIFDAGVAGVDPDIAGELP 483

419 DFGYDEDEHHYVDRKELIKYGSQVAPALEEILKNPCIDAVAVGIPDLAEGELP 478

484 GAVVVLKKGSMTEKEMVDYVAVSOVSNAKRLRGVRFVDEVPKGLTGKIDGKAIRELK 543

479 SAFVVKOPGKEITAKYDYLAERVSHTKYLRGCVRFVDSIPRVVTGKITRK---ELLKQ 535

544 PVAK 547

536 LLEK 539

RESULT 8

Photinus-luciferin 4-monoxygenase (ATP-hydrolyzing) (EC 1.13.12.7) OR [validated] -

N:Alternate names: firefly-type luciferase

C:Species: Pyrophorus plagiophthalmus

C:Date: 22-Nov-1993 #sequence-revision 03-Nov-1995 #text-change 02-Jun-2000

C:Accession: S29355

R:Wood, K.V.; Lam, Y.A.; Seliger, H.H.; McElroy, W.D.

Science 244, 700-702, 1989

A:Reference number: Z15086; MUID:96416441
A:Accession: T03789
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-542 <LEP>
A:Cross-references: EMBL:U50846; NID:g1663723; PIDN:AAB18638.1; PID:g1663724
C:Genetics:
A:Gene: ACL2
C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology
C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis
F:73-532/Domain: acetate-CoA ligase homology <ACL>

Query Match 26.5%; Score 748.5; DB 2; length 542;
Best Local Similarity 35.1%; Pred. No. 1.6e-45;
Matches 190; Conservative 89; Mismatches 200; Indels 63; Gaps 12;

QY 32 YMDRYAKGAIATN-----ALTGVD---YTYAEILEKSCCLGEALKNYGLVVDGR 79
DB 20 YIPNHLPLHSHYCEENISESSRPCLINGANKQIYTADELNSRKVAAGLHKGQIQPKDT 79
QY 80 IALSCENCEFFIPLVAGLFIQVAPFTNEITYLRELHSLGISKPTIYFSSKGLDKVI 139
DB 80 IMILFPNPEFVAFIAGSYLAI STMANPLFPAEVVQAKASSAKIITYQACHVANKV- 138
QY 140 TVQKTYTAKTIYILDSKYDYGQSMDFIK---KNTPOGPKGSSFKT-----V 186
DB 139 -----KDY---AFENDVKIICIDSAPEGCLHFSVLTQANEHDIPEV 176
QY 187 EYNRKEQVALIMNSSGSGTGLPKGVOLTHENLVTRFSHARDPIYGN-QVSPGTAILTVPE 245
DB 177 ELQPDQVVAL-PYSSGTGLPKGVMLTHGLVTSVAQVQDGENPNLYIHSEVMLCVLPL 235
QY 246 HHGFGMETTLGLVLTGER---IVMLTKFDETEFLTKDQKCSYILVPTLFAILNRSE 301
DB 236 FH---IYSLNSVLCGLRVGAAILIMQKFDIVSELELIQRYKVTIGPFVPIVLAIAKSP 292
QY 302 LLDKTDLSMLVEIASGAPLSKEIGAVARFRLPGVROGYGLTETTSAILI-----TPE 356
DB 293 MDDYDLSVSVKVMGSAAPLGELEDTVRAKFPNNAKLGQGYGTEAGPYLAMCLAFAPKEP 352
QY 357 GDDKPGASGVVPLFAKAVIDLTKKTLGPNRGEYCVKPMIMKGYVNDPEATREIIDE 416
DB 353 FEIKSGAGCTVVRNEMKIVDPKGTGSLPRNGSEICIRGDIIMKGYLNDPEATRTATIDK 412
QY 417 EGMULTGDIQYDEKHFIVRLKSLIKYKGYQVPALESVLLQHPNIFDAGVAGVPD 476
DB 413 EGMULTGDIQYIDDDDELIVRLKELIKYKGFVAPAELEALLNHPNISDAAVVPMKD 472
QY 477 PIAGELPGAVVVLKKGKSMTEKEVMDYVASOVSNAKRLGVRFRFDEVKGLTKIDGKA 536
DB 473 EQAGEVPVAFVVRNSGSTITTEDEVKDFISKQVITYFKRIK-RVFFVDALPKSPSGKILRKD 531
QY 537 IR 538
DB 532 LR 533

Search completed: November 29, 2001, 01:57:51
Job time: 361 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:57:51 ; Search time 36.66 Seconds
(without alignments)
1138.670 Million cell updates/sec

Title: US-09-581-241-6

Perfect score: 2823
Sequence: 1 MNNMNDENIVYGPPEPPYPI.....TGKIDKAIKREILKKPYAKM 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR.68:*
1: pit1:*
2: pit2:*
3: pit3:*
4: pit4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2814	99.7	548	1 S23437	Photinus-luciferin
2	2684	95.1	548	1 JS0181	Photinus-luciferin
3	2335.5	82.7	548	1 S33788	Photinus-luciferin
4	1967.5	69.7	547	2 S62787	Photinus-luciferin
5	1945.5	68.9	550	1 A26772	Photinus-luciferin
6	1367	48.4	543	2 S29354	Photinus-luciferin
7	1367	48.4	543	2 S29353	Photinus-luciferin
8	1363	48.3	543	2 S29355	Photinus-luciferin
9	1356	48.0	543	2 S29352	Photinus-luciferin
10	766.5	27.2	544	1 S01667	4-coumarate--COA 1
11	758.5	26.9	544	2 S15695	4-coumarate--COA 1
12	756	26.8	557	2 T07909	4-coumarate--COA 1
13	754	26.7	544	2 H85064	4-coumarate--COA 1
14	753.5	26.7	548	2 T07908	4-coumarate--COA 1
15	746.5	26.4	542	2 T03789	4-coumarate--COA 1
16	738.5	26.2	542	2 T02074	4-coumarate--COA 1
17	732.5	25.9	535	2 T08074	4-coumarate--COA 1
18	732.5	25.9	545	1 A39827	4-coumarate--COA 1
19	731	25.9	561	2 D96874	hypothetical prote
20	730	25.9	563	1 J00311	4-coumarate--COA 1
21	725.5	25.7	545	2 B39827	4-coumarate--COA 1
22	720	25.5	544	2 T20741	hypothetical prote
23	719	25.5	549	2 D86338	protein F5M15.18
24	718.5	25.5	561	2 S5784	4-coumarate--COA 1
25	713	25.3	570	2 T08075	4-coumarate--COA 1
26	691.5	24.5	569	2 T03390	4-coumarate--COA 1
27	680	24.1	537	2 T09755	4-coumarate--COA 1
28	674.5	23.9	542	2 B86654	hypothetical prote
29	672	23.8	537	2 T09710	4-coumarate--COA 1

ALIGNMENTS

30	659.5	23.4	423	2 PQ0772	4-coumarate--COA 1
31	635	22.5	552	2 E69438	probable fatty-aci
32	632.5	22.4	566	2 F85214	4-coumarate--COA 1
33	632.5	22.4	566	2 T05038	4-coumarate--COA 1
34	600	21.3	566	2 H84037	long-chain fatty-a
35	597.5	21.2	569	2 C69471	probable fatty-aci
36	594.5	21.1	564	2 G84037	long-chain fatty-a
37	587.5	20.8	293	2 S31705	4-coumarate--COA 1
38	581.5	20.6	593	2 E69378	probable acid--COA
39	577	20.4	548	2 D69187	probable acid--COA
40	574.5	20.4	560	2 D69649	probable long-cha
41	573	20.3	540	2 T18841	hypothetical prote
42	568	20.1	561	2 S41589	long-chain-fatty-a
43	568	20.1	561	2 E85791	hypothetical prote
44	565	20.0	546	2 C83791	acid-CoA ligase BH
45	564.5	20.0	584	2 C75364	probable long-cha

RESULT 1	
S23437	Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - Japanese firefly
N:Alternate names:	firefly luciferase
C:Species:	Luciola lateralis (Japanese firefly)
C:Date:	22-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 02-Jun-2000
C:Accession:	S23437; S57417
R:Ratsumi, H.; Kajiyama, N.; Nakano, E.	
Biochim. Biophys. Acta 1131, 161-165, 1992	
A:Title:	Molecular cloning and expression in Escherichia coli of a cDNA clone encodin
A:Reference number:	S23437; MUID:92305054
A:Accession:	S23437
A:Molecule type:	mRNA
A:Residues:	1-548 <TRAIL>
A:Cross-references:	EMBL:X66919; NID:99526; PIDN:CAA47358.1; PID:99527
R:Cho, K.; Choi, Y.; Boo, K.	
submitted to the EMBL Data Library, June 1995	
A:Description:	Molecular cloning of gene for luciferase in Luciola lateralis.
A:Reference number:	S57417
A:Accession:	S57417
A:Molecule type:	DNM
A:Residues:	1-4,'D',6-14,'K',16-145,'N',147-175,'P',177-505,'G',507-548 <CHO>
A:Cross-references:	EMBL:Z49991; NID:9871400; PIDN:CAA90072.1; PID:9871401
C:Genetics:	
A:introns:	44/3; 114/1; 225/2; 332/2; 450/3; 506/1
C:Function:	
A:Description:	catalyzes the irreversible reaction of luciferin with oxygen and ATP t
A:Note:	magnesium is required; color of emitted light varies among species from green
C:Superfamily:	4-coumarate--COA ligase; acetate--COA ligase homology
C:Keywords:	ATP; luminescence; magnesium; monooxygenase; oxidoreductase; peroxisome
F:73-537/Domain:	acetate--COA ligase homology <ACI>

Query Match	99.7%	Score 2814;	DB 1;	Length 548;
Best Local Similarity	99.6%	Pred. No. 2.7e-192;		
Matches 546;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	MNNMNDENIVYGPPEPPYPIEEGSAGAOQLKRYMDRAKIGALIAFTNALGVDTYAYLE 60
DB	1	MNNMNDENIVYGPPEPPYPIEEGSAGAOQLKRYMDRAKIGALIAFTNALGVDTYAYLE 60
QY	61	KSCCLGALKNYGLVVDRIALCSNCEEFIVLAGLFGVAPPTNEIYTLRELVSHL 120
DB	61	KSCCLGALKNYGLVVDRIALCSNCEEFIVLAGLFGVAPPTNEIYTLRELVSHL 120
QY	121	GISKPTVFSKKGDKLVITVOKTVAIKTIVILDSVDVDRGVQSDMNFKKTPPGCFG 180
DB	121	GISKPTVFSKKGDKLVITVOKTVAIKTIVILDSVDVDRGVQSDMNFKKTPPGCFG 180
QY	181	SSFTVENVNRKEQVALIMNSGSGTGLPKGVOLTHENIVTRFSHARDPIGVNOVSPGTAIL 240
DB	181	SSFTVENVNRKEQVALIMNSGSGTGLPKGVOLTHENIVTRFSHARDPIGVNOVSPGTAIL 240


```

QY 241 TVVPHHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTLDYKCSSVILVPTLAILNRS 300
    |||
Db 241 TVVPHHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTLDYKCSSVILVPTLAILNRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSEIGEAARFNLPVNGVGLTETTSAILITPBGDK 360
    |||
Db 301 ELLDKYDLSNLVEIASGAPLSEIGEAARFNLPVNGVGLTETTSAILITPBGDK 360
QY 361 PGASGKVVPLFAKAVIDLDTKKTLPGRNGEVCVKPMLMKGVNDPEATRETIIDEGWL 420
    |||
Db 361 PGASGKVVPLFAKAVIDLDTKKTLPGRNGEVCVKPMLMKGVNDPEATRETIIDEGWL 420
QY 421 HTGDIGYDEEKHFIVDRKLSLKIKYGYVPPAELESVLLQHPNIFDAGVAGVPPDIPG 480
    |||
Db 421 HTGDIGYDEEKHFIVDRKLSLKIKYGYVPPAELESVLLQHPNIFDAGVAGVPPDIPG 480
QY 481 ELPGAVVVLKKGKSMTEKEMDVASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIRI 540
    |||
Db 481 ELPGAVVVLKKGKSMTEKEMDVASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIRI 540
QY 541 LKKPVAKM 548
    |||
Db 541 LKKPVAKM 548

```

RESULT 2

```

Photinus-luciferin 4-monoxygenase (Atp-hydrolyzing) (EC 1.13.12.7) - GenJ1 firefly
N:Alternate names: firefly luciferase
C:Species: Luciola cruciata (GenJ1 firefly)
C>Date: 07-Jun-1990 #sequence_revision 20-Aug-1994 #text_change 26-May-2000
C:Accession: JS0181
R:Masuda, T.; Tatsumi, H.; Nakano, E.
Gene 77, 265-270, 1989
A:Title: Cloning and sequence analysis of cDNA for luciferase of a Japanese firefly, Luc
A:Reference number: JS0181; MUID:89326143
A:Accession: JS0181
A:Molecule type: mRNA
A:Residues: 1-548 <MAS>
A:Cross-references: GB:M26194; NID:g159050; PIDN:AAA29135.1; PID:g159051
C:Comment: This protein catalyzes the oxidation of luciferin in the presence of ATP, oxy
    pecies from green to yellow.
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: Atp; luminescence; monoxygenase; oxidoreductase
F:73-537/Domains: acetate--CoA ligase homology <ACLD>

```

```

Query Match 95.1%; Score 2684; DB 1; Length 548;
Best Local Similarity 93.4%; Pred. No. 4,7e-183;
Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

```

```

QY 1 MENNENENIYGGPEPPIIEGSAQOLRKYMRYAKLGAIAFTNALTGVDYTAYLE 60
    |||
Db 1 MENNENENIYGGPEPPIIEGSAQOLRKYMRYAKLGAIAFTNALTGVDYTAYLE 60
QY 61 KSCGLALKNYGLVNDGRIALSCENCEEFPPIVLAGLFGVGAAPNEIYTLRELVHSL 120
    |||
Db 61 KSCGLALKNYGLVNDGRIALSCENCEEFPPIVLAGLFGVGAAPNEIYTLRELVHSL 120
QY 121 GISKPTLVSSKGLDVIYVOKTVTAIKTIVILDSKVDYRGYSMDNFKKNTPOGFKG 180
    |||
Db 121 GISKPTLVSSKGLDVIYVOKTVTAIKTIVILDSKVDYRGYSMDNFKKNTPOGFKG 180
QY 181 SSRKYTVENRKEOVALIMNSSGSTGLPKGVOLNHNITVTRSHARDPITYGNOVSPGTAVL 240
    |||
Db 181 SSRKYTVENRKEOVALIMNSSGSTGLPKGVOLNHNITVTRSHARDPITYGNOVSPGTAVL 240
QY 241 TVVPHHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTLDYKCSSVILVPTLAILNRS 300
    |||
Db 241 TVVPHHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTLDYKCSSVILVPTLAILNRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSEIGEAARFNLPVNGVGLTETTSAILITPBGDK 360
    |||
Db 301 ELLDKYDLSNLVEIASGAPLSEIGEAARFNLPVNGVGLTETTSAILITPBGDK 360

```

```

Db 301 ELLDKYDLSNLVEIASGAPLSEIGEAARFNLPVNGVGLTETTSAILITPBGDK 360
    |||
QY 361 PGASGKVVPLFAKAVIDLDTKKTLPGRNGEVCVKPMLMKGVNDPEATRETIIDEGWL 420
    |||
Db 361 PGASGKVVPLFAKAVIDLDTKKTLPGRNGEVCVKPMLMKGVNDPEATRETIIDEGWL 420
QY 421 HTGDIGYDEEKHFIVDRKLSLKIKYGYVPPAELESVLLQHPNIFDAGVAGVPPDIPG 480
    |||
Db 421 HTGDIGYDEEKHFIVDRKLSLKIKYGYVPPAELESVLLQHPNIFDAGVAGVPPDIPG 480
QY 481 ELPGAVVVLKKGKSMTEKEMDVASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIRI 540
    |||
Db 481 ELPGAVVVLKKGKSMTEKEMDVASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIRI 540
QY 541 LKKPVAKM 548
    |||
Db 541 LKKPVAKM 548

```

RESULT 3

```

Photinus-luciferin 4-monoxygenase (Atp-hydrolyzing) (EC 1.13.12.7) [validated] - sou
S33788
N:Alternate names: firefly luciferase
C:Species: Luciola mingrellica (southern Russian firefly)
C>Date: 02-Dec-1993 #sequence_revision 03-Nov-1995 #text_change 02-Jun-2000
C:Accession: S33788
R:Devine, J.H.; Kutuzova, G.D.; Green, V.A.; Ugarova, N.N.; Baldwin, T.O.
Biochim. Biophys. Acta 1173, 121-132, 1993
A:Title: Luciferase from the East European firefly Luciola mingrellica: cloning and nu
A:Reference number: S33788; MUID:93277944
A:Accession: S33788
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-548 <DEV>
A:Cross-references: GB:S61961; NID:g409316; PIDN:AAB26932.1; PID:g409317
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: Atp; luminescence; monoxygenase; oxidoreductase
F:72-537/Domains: acetate--CoA ligase homology <ACLD>

```

```

Query Match 82.7%; Score 2335.5; DB 2; Length 548;
Best Local Similarity 81.7%; Pred. No. 2.8e-158;
Matches 446; Conservative 47; Mismatches 52; Indels 1; Gaps 1;

```

```

QY 4 MENNENIYGGPEPPIIEGSAQOLRKYMRYAKLGAIAFTNALTGVDYTAYLESC 63
    |||
Db 3 MEKENVYVGLPFRPIIEGSAQOLRKYMRYAKLGAIAFTNALTGVDYTAILESC 62
QY 64 CLGEALKNYGLVNDGRIALSCENCEEFPPIVLAGLFGVGAAPNEIYTLRELVHSLGIA 123
    |||
Db 63 RLAEAMKNFGKPREHIALSCENCEEFPPIVLAGLFGVGAAPNEIYTLRELVHSLGIA 122
QY 124 KPTIVSSKGLDVIYVOKTVTAIKTIVILDSKVDYRGYSMDNFKKNTPOGFKGSSF 183
    |||
Db 123 OPTIVSSKGLDVIYVOKTVTAIKTIVILDSKVDYRGYSMDNFKKNTPOGFKGSSF 182
QY 184 KTVEN-NRKEOVALIMNSSGSTGLPKGVOLNHNITVTRSHARDPITYGNOVSPGTAVL 242
    |||
Db 183 VPIDVKNRKHVALIMNSSGSTGLPKGVOLNHNITVTRSHARDPITYGNOVSPGTAVL 242
QY 243 VPHHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTLDYKCSSVILVPTLAILNSEL 302
    |||
Db 243 VPHHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTLDYKCSSVILVPTLAILNSEL 302
QY 303 LDKYDLSNLVEIASGAPLSEIGEAARFNLPVNGVGLTETTSAILITPBGDKPG 362
    |||
Db 303 LDKYDLSNLVEIASGAPLSEIGEAARFNLPVNGVGLTETTSAILITPBGDKPG 362
QY 363 ASGKVVPLFAKAVIDLDTKKTLPGRNGEVCVKPMLMKGVNDPEATRETIIDEGWLHT 422
    |||
Db 363 ASGKVVPLFAKAVIDLDTKKTLPGRNGEVCVKPMLMKGVNDPEATRETIIDEGWLHT 422

```


QY 423 GDIGYDEEHFFIVDLKSLIKYGYVPPALESVLLQHPNIFDAGVAGVPPDPAIGEL 482
 |||||
 Db 423 GDIGYDEEHFFIVDLKSLIKYGYVPPALESVLLQHPNIFDAGVAGVPPDPAIGEL 482
 QY 483 PGAVVVLKSKSMTEKEVMDYVAVSOVSNARKLRGVAFVDEVEKGLTGKIDGKAIREILK 542
 |||||
 Db 483 PGAVVVLKSKSMTEKEVMDYVAVSOVSNARKLRGVAFVDEVEKGLTGKIDGKAIREILK 542
 QY 543 KPVAKM 548
 |||||
 Db 543 KPVAKM 548

RESULT 4
 S62787
 Photinus-luciferin 4-monoxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - Lampyris noctiluca
 N:Alternate names: firefly luciferase
 C:Species: lampyris noctiluca
 C>Date: 23-Aug-1996 #sequence_revision 01-Nov-1996 #text_change 26-May-2000
 C:Accession: S62787; S57949
 R:Salal-Newby, G.B.; Thomson, C.M.; Campbell, A.K.
 Biochem. J. 313, 761-767, 1996
 A:Title: Sequence and biochemical similarities between the luciferases of the glow-worm
 A:Reference number: S62787; MUID:96190714
 A:Accession: S62787
 A:Molecule type: mRNA
 A:Residues: 1-547 <SAL>
 A:Cross-references: EMBL:X89479; NID:g989314; PIDN:CAA61668.1; PID:g989315
 C:Superfamily: 4-comumate--CoA ligase; acetate--CoA ligase homology
 C:Keywords: ATP; luminescence; monoxygenase; oxidoreductase
 F:71-535/Domain: acetate--CoA ligase homology <ACL>

Query Match 69.7%; Score 1967.5; DB 2; Length 547;
 Best Local Similarity 69.0%; Pred. No. 4.1e-132;
 Matches 372; Conservative 76; Mismatches 90; Indels 1; Gaps 1;

QY 4 MENDENIVYGEPEFYPIEESAGAQLRKYMDRYAKL-GATAFTNALTGVDYTAETLEKS 62
 |||||
 Db 1 MEDAKNIMKGRAPFYLEDGTAGEQLHKMKRKYVPGTIAFTNAHEVNITTYEYFEMA 60
 QY 63 CCLGALKNYGLVVDGRIALCSECEFFIPVLAGLFTGVGAPDPTNETYTLRELVHSLGI 122
 |||||
 Db 61 CRLAETMKRYGLGLOHNIHIVCSENSLOPFMPVCGALFVGAVASTNDIYNRELINSLSI 120
 QY 123 SKPTIVSSKKGLDKVITVQKTAIKTIYVLDKVDYRGQSDNFIKKNTPGFGKSS 182
 |||||
 Db 121 SOPTIVSSSKRALOKILGVOKRPLIKRIYVLDSDREDYMGQSMVSELSHLPAGFNEYD 180
 QY 183 FRTVEVNRKEDVALIMNMSGSTGLPKGVOLTHENIVTRFESHARDPIYGNQVSPGTAITLV 242
 |||||
 Db 181 YIPDSFDEETATALIMNMSGSTGLPKGVOLTHENIVTRFESHARDPIYGNQVSPGTAITLV 240
 QY 243 VPHHGFMTTGLGYLTCGRFVLMFKFDEFTFLKTDYKCSVILVPTLFAILINSEL 302
 |||||
 Db 241 IPEFHGFMFTTGLGYLTCGRFVLMFKFDEFTFLKTDYKCSVILVPTLFAILINSEL 300
 QY 303 LDKYVLSNLVEIASGAPLAKSEIGAVARRNLPVRCGYGLTETTSALITTPGDDKPG 362
 |||||
 Db 301 VDKYVLSNLVEIASGAPLAKSEIGAVARRNLPVRCGYGLTETTSALITTPGDDKPG 360
 QY 363 ASGVVPLFKAKVYIDLTKTTLGPNRRGEVCKGPMILKKGVDNPEAREIIDEESWLHT 422
 |||||
 Db 361 AGCKVVPFSAKIYDLDGTGLGVQNGELCKGPMIKGYVNNPEARISALIDKGMWLS 420
 QY 423 GDIGYDEEHFFIVDLKSLIKYGYVPPALESVLLQHPNIFDAGVAGVPPDPAIGEL 482
 |||||
 Db 421 GDIAVYDKDGHFFIVDLKSLIKYGYVPPALESVLLQHPNIFDAGVAGVPPDPAIGEL 480
 QY 483 PGAVVVLKSKSMTEKEVMDYVAVSOVSNARKLRGVAFVDEVEKGLTGKIDGKAIREIL 541
 |||||
 Db 481 PAAVVVLKSKSMTEKEVMDYVAVSOVSNARKLRGVAFVDEVEKGLTGKIDGKAIREIL 539

RESULT 5
 A26772
 Photinus-luciferin 4-monoxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - common eastern
 N:Alternate names: firefly luciferase
 C:Species: Photinus pyralis (common eastern firefly)
 C>Date: 19-Nov-1988 #sequence_revision 20-Aug-1994 #text_change 26-May-2000
 C:Accession: A26772
 R:de Wet, U.R.; Wood, K.V.; Deluca, M.; Hellinski, D.R.; Subramani, S.
 Mol. Cell. Biol. 7, 725-737, 1987
 A:Title: Firefly luciferase gene: structure and expression in mammalian cells.
 A:Reference number: A26772; MUID:87144243
 A:Accession: A26772
 A:Molecule type: DNA
 A:Residues: 1-550 <DEW>
 A:Cross-references: GB:M15077; NID:g160793; PIDN:AAA29795.1; PID:g160794
 A:Note: The authors translated the codon CAA for residue 134 as Glu
 C:Comment: This protein catalyzes the oxidation of luciferin in the presence of ATP,
 C:Genetics: pieces from green to yellow.
 A:introns: 42/3; 112/1; 223/2; 330/2; 448/3; 504/1
 C:Superfamily: 4-comumate--CoA ligase; acetate--CoA ligase homology
 C:Keywords: ATP; luminescence; monoxygenase; oxidoreductase
 F:71-535/Domain: acetate--CoA ligase homology <ACL>
 F:548-550/Region: peroxisome/glyoxysome location signal (S-[RKH]-L) motif

Query Match 68.9%; Score 1945.5; DB 1; Length 550;
 Best Local Similarity 68.2%; Pred. No. 1.5e-130;
 Matches 369; Conservative 75; Mismatches 96; Indels 1; Gaps 1;

QY 4 MENDENIVYGEPEFYPIEESAGAQLRKYMDRYAKL-GATAFTNALTGVDYTAETLEKS 62
 |||||
 Db 1 MEDAKNIMKGRAPFYLEDGTAGEQLHKMKRKYVPGTIAFTNAHEVNITTYEYFEMA 60
 QY 63 CCLGALKNYGLVVDGRIALCSECEFFIPVLAGLFTGVGAPDPTNETYTLRELVHSLGI 122
 |||||
 Db 61 VRLAEMKRYGLTNTNRIVYVCSNSLOPFMPVCGALFVGAVAPADIDYNEBELLSMTI 120
 QY 123 SKPTIVSSKKGLDKVITVQKTAIKTIYVLDKVDYRGQSDNFIKKNTPGFGKSS 182
 |||||
 Db 121 SOPTIVSSSKRALOKILGVOKRPLIKRIYVLDSDREDYMGQSMVSELSHLPAGFNEYD 180
 QY 183 FRTVEVNRKEDVALIMNMSGSTGLPKGVOLTHENIVTRFESHARDPIYGNQVSPGTAITLV 242
 |||||
 Db 181 FVPEFDRDKTIALIMNMSGSTGLPKGVOLTHENIVTRFESHARDPIYGNQVSPGTAITLV 240
 QY 243 VPHHGFMTTGLGYLTCGRFVLMFKFDEFTFLKTDYKCSVILVPTLFAILINSEL 302
 |||||
 Db 241 VPHHGFMTTGLGYLTCGRFVLMFKFDEFTFLKTDYKCSVILVPTLFAILINSEL 300
 QY 303 LDKYVLSNLVEIASGAPLAKSEIGAVARRNLPVRCGYGLTETTSALITTPGDDKPG 362
 |||||
 Db 301 IDKYVLSNLVEIASGAPLAKSEIGAVARRNLPVRCGYGLTETTSALITTPGDDKPG 360
 QY 363 ASGVVPLFKAKVYIDLTKTTLGPNRRGEVCKGPMILKKGVDNPEAREIIDEESWLHT 422
 |||||
 Db 361 AVGVVVPFSAKIYDLDGTGLGVQNGELCKGPMIKGYVNNPEARISALIDKGMWLS 420
 QY 423 GDIGYDEEHFFIVDLKSLIKYGYVPPALESVLLQHPNIFDAGVAGVPPDPAIGEL 482
 |||||
 Db 421 GDIAVYDEEHFFIVDLKSLIKYGYVPPALESVLLQHPNIFDAGVAGVPPDPAIGEL 480
 QY 483 PGAVVVLKSKSMTEKEVMDYVAVSOVSNARKLRGVAFVDEVEKGLTGKIDGKAIREIL 542
 |||||
 Db 481 PAAVVVLKSKSMTEKEVMDYVAVSOVSNARKLRGVAFVDEVEKGLTGKIDGKAIREIL 540
 QY 543 K 543
 |||||
 Db 541 K 541

RESULT 6

S29354
Phospho-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) YE [validated] - lum
N.Alternate names: firefly-type luciferase
C.Species: Pyrophorus plagiophthalmus
C.Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text_change 02-Jun-2000
R.Accession: S29354
R.Wood, K.V.; Lam, Y.A.; Seliger, H.H.; McElroy, W.D.
Science 244, 700-702, 1989
A.Title: Complementary DNA coding for firefly luciferases can elicit bioluminescence of
A.Reference number: S29352; MUID:89242142
A.Accession: S29354
A.Status: Preliminary
A.Molecule type: mRNA
A.Residues: 1-543 <WOO>
C.Superfamily: 4-connarate--CoA ligase: acetate--CoA ligase homology
C.Keywords: ATP: luminescence; monooxygenase; oxidoreductase
F:70-532/Domain: acetate--CoA ligase homology <ACL>

Query Match	48.4%;	Score 1367;	DB 2;	Length 543;
Best Local Similarity	48.3%;	Pred. No. 2e-89;		
Matches 263;	Conservative 108;	Mismatches 167;	Indels 6;	Gaps 3

[illegible]

A>Title: Complementary DNA coding click beetle luciferases can elicit bioluminescence
A:Reference number: S29352; MUID:85242142
A:Accession: S29352
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-543 <MO>
C:Superfamily: 4-conaminate-CoA ligase; acetate--CoA ligase homology
C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase
F:70-532/Domain: acetate--CoA ligase homology <ACL>

Query Match 48.4%; Score 1367; DB 2; Length 543;
Best Local Similarity 49.1%; Pred. No. 2e-89;
Matches 267; Conservative 103; Mismatches 168; Indels 6; Gaps 3;

[illegible]

RESULT 7
S29353
Phytinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) YC [validated] - 1um
N:Alternate names: firefly-type luciferase
C:Species: Pyrophorus plagiophthalmus
C:Date: 25-Feb-1994 #sequence-revision 01-Sep-1995 #text-change 02-Jun-2000
C:Accession: S29353
R:Wood, K.V.; Lam, Y.A.; Seliger, H.H.; McElroy, W.D.
Science 244, 700-702, 1989

RESULT 8
S29355
Photinus-luciferin 4-monoxygenase (ATP-hydrolyzing) (EC 1.13.12.7) OR [validated] -
N:Alternate names: firefly-type luciferase
C:Species: Pyrophorus plagiophthalmus
C:Date: 22-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 02-Jun-2000
C:Accession: S29355
R:Wood, K.V.; Lam, Y.A.; Seliger, H.H.; McElroy, W.D.
Science 244, 700-702, 1989
A:Title: Complementary DNA coding for beetle luciferases can elicit bioluminescence
A:Reference number: S29352; MVID:89242142
A:Accession: S29355
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-543 <MO>
C:Superfamily: 4-coumate--CoA ligase; acetate--CoA ligase homology
:Keywords: ATP; luminescence; monooxygenase; oxidoreductase

F;70-532/Domain: acetate--CoA ligase homology <ACL>

Query Match	48.3%;	Score 1363;	DB 2;	Length 543;
Best Local Similarity	48.3%;	Pred. No. 3.9e-89;		
Matches 253; Conservative	108;	Mismatches 167;	Indels 6;	Gaps 3

[illegible]

Query Match	48.0%	Score 1356	DB 2	Length 543
Best Local Similarity	48.4%	Pred. No. 1.2e-88		
Matches 265	Conservative 105	Mismatches 164	Indels 14	Gaps 4
0y	4	MENDENIVGPEPPYPTPEESGAGQLRKWDRAKGAIAFTNALGVDP----	YTYAEYL	59

Db 2 MKRKNNVYGPRLPHLEDLTAGMLPRLARKSHL-----PQALVDVYGEEWISKEEF 56

QY 60 EKSCCLGALKNGVLVVDGRIALCSENCSEEFPIPVLAGLFGVGAVPMEITTLRELVHS 119

Db 57 EFTCLLAOSLHNCGYCKKMSDVYSICAENNRKRFEPDIIAAMYGIVAVNEGYPIDELCRV 116

QY 120 LGISKPTVYESSKKGLDKVITVQKTVTAIKTIYILDSKVYRGVQSMDDIKKNTPOGFR 179

Db 117 MGISRPLVQVETKNIKLNKLVLEVOSTRTDFIKRIIILDVENIHGCESLPNFTSKYSDGNI- 175

QY 180 GSFETVEVNRKEQVALIMNSGSGTLPGVOLTENIIVTRFSHARDPIYGNVSEGTAI 239

Db 176 -ANKRPLHYDVEQVAAILCSGFTGLPKGMQTNHNVRCVRLIHALDPVGTQLIGVTY 234

QY 240 LTVVPFHFGMFTTGLTGTCGFRIVMLTKRDEETFLKTIQDYKCSSVILVPLFLALNR 299

Db 235 LVYLPFFHAFEFISNLGTFWGLRVIMLRPDQAFPKAIQADYEVRSVINVPAILFLSK 294

QY 300 SELLDKYLDSNLVETIASGAPLSEIKGEIVAVRRNLPGRCOGVGLTFTTSAIITTEGSD 359

Db 295 SPLVDKYLDSLRELCGCAAPLAKAEVIAVKRNLEIGIKCGRLTSESTANIHSLRDER 354

QY 360 KFGAGSKVVPLEKAKVIDLDTRKTLGPNRNGEVCVKPMLKGVNDPEARTREIIDEGW 419

Db 355 KGSGLGRVTPLEMAKIAIDRETGKALGPNQVGLCTIKGPMVSKGYVNNVEATKEAIDDDGW 414

QY 420 IATGDTGYDEKHEFFIYDLRLSLIKRYGYVPAPELESYILQHPNIFDAGVAGVPPPIA 479

Db 415 IHSFGGYDDEBHHYVYDRYKELIKYKGSVAPAELEILLKPPCIRDAVAVGIPDLEA 474

QY 480 GELPAAVVVLKKGKSMTEKEVMYDVASOVSAKRLRGSVRFVDVPKGLTGIDKAIKE 539

Db 475 GELPSAFVYIQGKEITTAKEYVDIYLAERVSHTKYLRGSGVRVDSIPRNVTGKITRK--E 531

QY 540 ILKKPVAK 547

Db 532 LTKLLEK 539

RESULT 10

S01667

4-commarate--CoA ligase (EC 6.2.1.12) (clone pc4CL-1) - parsley

N:Alternate names: 4-commaroyl-CoA synthetase

C:Species: Petroselinum crispum (parsley)

C:date: 30-Sep-1989 #sequence_revision 20-Aug-1994 #text_change 05-May-2000

R:Accession: S01667

R:Lozoya, E.; Hofmann, H.; Douglas, C.; Schulz, W.; Scheel, D.; Hahlbrock, K. Eur. J. Biochem. 176, 661-667, 1988

A:Title: Primary structures and catalytic properties of isoenzymes encoded by the two A:Reference number: S01667; MUID:89005119

A:Accession: S01667

A:Molecule type: mRNA

A:Residues: 1-544 <L00Z>

A:Cross-references: EMBL:X13324; NID:g20431; PIDN:CAJ1696.1; PID:g20432

A:Gene: 4CL-1

C:Superfamily: 4-commarate--CoA ligase; acetate-CoA ligase homology

C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis

F:74-333/Domain: acetate-CoA ligase homology <ACL>

Query Match	27.2%;	Score 766.5;	DB 1;	Length 544;
Best Local Similarity	36.0%;	Pred. No. 1e-46;		
Matches 186;	Conservative	89;	Mismatches 189;	Indels 53; Gaps 10;

Qy	46	NATLGVDYTAFAELKSCCLG	EALKNYGLVVDGRIM	ALCSENEEEFFIPYLA	GLFTGVGA	105
Db	47	NCATGEFFITSQVELLSKRV	ASGLNKLGIOQDITML	LPNSPEYFPFLAS	SYNGAIST	106
Qy	106	PTNEIYTLRELVLHSLG	SLFTIVSSSKGLD	KVITVQKTVAKIT	IVL--DSK	VDYRG 162
Db	107	MANPFTSASVIVIKQLAS	QAKLIITQAC	YVDK---KDYAEK	NIQIICID	DARQDCLH 162

```

RESULT 11
S15695
4-comumarate--CoA ligase (EC 6.2.1.12) (clone pcdCL-2) - parsley
C:Species: Petroselinum crispum (parsley)
C:Date: 13-Jan1995 #sequence_revision 13-Jan-1995 #text_change 05-May-2000
C:Accession: S15695
R:Lozoya, E.; Hoffmann, H.; Douglas, C.; Schultz, W.; Scheel, D.; Hahlbrock, K.
Eur. J. Biochem. 176, 661-667, 1988
A:Title: Primary structures and catalytic properties of isoenzymes encoded by the two 4
A:Reference number: S01667; MUID:89005119
A:Accession: S15695
A:Molecule type: mRNA
A:Residues: 1-544 <L02>
A:Cross-references: EMBL:X13325; NID:g20435; PIDD:CAA31697.1; PID:g20436
C:Genetics:
A:Gene: 4CL-2
C:Superfamily: 4-comumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: acid-thiol ligase; coenzyme A
F:74-533/Domain: acetate--CoA ligase homology <ACL>

```

Db 259 MOKRDIIVPLELLELQKKVITIGFEVPIVLAIAKSPVADKIDLSVTRTWMSGAAPIGKETE 318

QY 327 EAVAREFNLPVROGQGLTETTSATII-----TPBGDDKPGASGVVPLFAKAVYIDLDTK 381

Db 319 DAVRAKFPNAKIGQIGMTAGCPVLAMCLAFAKEPEIKSGAGTVVRAEKKIYDPETN 378

QY 382 KTLGNRRGCEVCVGMPLMKKGVNDPEATRETIDEGMHTGDIGYDEEKKHFFIVDRK 441

Db 379 ASLPRNORGEICIRGDIQKMGYLANDPESTRITIDEGMHTGDIGFIDDDDELFIYDLK 438

QY 442 SLIKKGYVPAPELESTYLQHPNIFDAGVAVPPPIAGELPGAAVYLKKGSKMTEKVM 501

Db 439 ELIKIKGVPAPELEALLLHPITISDAVAPMIDEKAGEVPAVAVRTNPTTEEBIK 498

QY 502 DYVASOYNAKRLRGVRFVDEVPKGLTGKIDGKAIR 538

Db 499 QFVSKOVVEFYKRI-PRFEVDAIPKPSGKILRKDLR 534

RESULT 12

707909

4:comarate--CoA ligase (EC 6.2.1.12) 1 [validated] - western balsam poplar x cottonwood
C:Species: Populus trichocarpa x Populus deltoides (western balsam poplar x cottonwood)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 26-May-2000
C:Accession: T07909

R.Allina, S.M.; P.I.-Nadash, A.; Theilmann, D.A.; Ellis, B.E.; Douglas, C.J.
Plant Physiol. 116, 743-754, 1998

A:Title: 4-Comarate:Coenzyme A ligase in hybrid poplar. Properties of native enzymes
A:Reference number: Z16208; MUID:98150279

A:Accession: T07909

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-557 <ALL>

A:Cross-references: EMBL:AF008184; NID:g2911798; PIDN:AC39366.1; PID:g2911799

C:Genetics:

A:Gene: ACCL

C:Function:

A:Description: EC 6.2.1.12; provides activated thioester substrates for phenylpropano
A:Superfamily: 4-comarate--CoA ligase; acetate--CoA ligase homology
A:Keywords: acid-thiol ligase; coenzyme A; flavinoid biosynthesis
F;74-534/Domain: acetate--CoA ligase homology <ACCL>

Query Match 26.84; Score 756; DB 2; Length 557;
Best Local Similarity 35.58; Pred. No. 5,9e-46;
Matches 183; Conservative 84; Mismatches 199; Indels 50; Gaps 11;

QY 46 NALGVDTYAEYLEKSCCLGEALKNYGLVWDGRIALCSECEEFPIPLAGLFIGVA 105

Db 47 NGPGEIHTYAEVLETSRKVYASGLNKGIDVILLLONSPEFYAPAFGASITGIST 106

QY 106 PTNPIITLRELVHSLGSKPIYSSKKGIDKTYTVOKTYTAIKTYIILDSKVDRYOS 165

Db 107 TANFTYPAEVAQAATASAKLLITQAVVAEKVOEFKENVHAKIYVDSPPENYLFSE 166

QY 166 MDNFIKNTPOGEGSSFTVEVNRKEOVAALIMSSGSLPKGYOLTENHIYTR----- 220

Db 167 LTNSDEDIP-----AVEINPDYVAL-PPSSGTTILPRGWMILTKHGLVTSVAQOV 216

QY 221 -----FSNARDPIYONOVSPGTAILTVPRFHNGGMEFTTGLYLCGFR---IYMLTK 269

Db 217 DGENPMLYFHEKD-----VILCVLPFH-----IYSLNSVLLCGLRVGSAILLMOK 262

QY 270 FEDERTLKTQADKCSSVILPTLFAILNSELDDKYDLSNLEVIASGGAPSLKEIGAV 329

Db 263 FEIYITLMELOKRVKVIAPFPVPLAVAKCPYVDKGLDSSITVWGAAPKKELEDIV 322

QY 330 AKRFNLPQAV--QGYGITE---TSAILITPBG--DDKPGASGVVPLFAKAVYIDLDTK 382

Db 323 --RAKLPMALGGQYGTENGAPVLSMCLAFAPKPEPIKSGAGCTVVRNEMKIVDPDGR 380

QY 383 TLDPNRGEVCVGMPLMKGVVNDPEATRETIDEGMHTGDIGYDEEKKHFFIVDRKS 442

Db 381 SLPRNDSGEICIRGSGIMKGYLNDEPEATERVNDGMLHTTGIDIGYIDGDELFIYDRKE 440
 QY 443 LKRYGYVPAELESVLLQHNIFEDAGVAGVDPPIAGCLPGAVVVLKKGKSMTEKEVMD 502
 Db 441 LKRYGYVPAELESVLLQHNIFEDAGVAGVDPPIAGCLPGAVVVLKKGKSMTEKEVMD 500
 QY 503 YVASOVNAKRLRGVRFVDEVPKGLTGKIDKKAIR 538
 Db 501 YISKOVIFYKRI-SRFFTEALPKAPSGKILKDKLR 535

RESULT 13

H85064
 4-coumarate--CoA ligase-like protein (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
 C:Accession: H85064
 R:anonymus, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 Nature 402, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: AB5001; MUID:20083488
 A:Accession: H85064
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-544 <STO>
 A:Cross-references: GB:NC_001268; NID:g7267275; PIDN:CAB81058.1; GSPDB:GN00140
 C:Genetics:
 A:gene: AT4G05160
 A:Map position: 4
 C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

Query Match 26.7%; Score 754; DB 2; Length 544;
 Best Local Similarity 34.0%; Pred. No. 7.9e-46;
 Matches 178; Conservative 111; Mismatches 195; Indels 40; Gaps 12;

QY 42 IAFNALGVDTYAYEYLEKSCGLKNGVLDGRIALCSENCPEFFIYVLAGLFTG 101
 Db 44 LATADSDTGLTFSSQLKSAVARLHGFRLGRKNDVLLFAPNSYCPPLCEFLAVTAIG 103
 QY 102 VCAAPNNEYITRELVSIGISKPTIV-----FSSKKGLDKVITYQKIVTAIKTIYLD 155
 Db 104 GVFTTANPLYTYEVESKQIKDSNPKIIISVNOLEFKIKFDPVVLGSKDVEIPEPGSN 163
 QY 156 SKVDYAGYOSMDNFIKKNPOGKSSFKTYEVRKEQVALIMNSGSGTLPKQVOLTRE 215
 Db 164 SKI-----LSFDNVMELSPV-----SEYPRVELKQSDTALLY-SSGTTGSKGVELTHG 213
 QY 216 N-----IYTRSHADPIYGNQVSPGTAILLVPPFHGFGM-FTTLGYLTCGFRIVMLT 268
 Db 214 NFIAASLWMTMODDLNGEYHG-----VELCFLPWFHVFGLAVITYSQLQGNMLVSMA 266
 QY 269 KPEEFELTKLDYKCSSVILVPTLFAILNRSFLDKYDLSNIVETASGAPLSKETGA 328
 Db 267 REFLEVLVNLNIEKFRVTHLMVPPVFLALSKOSIYKFLDLSLKYSGAAPLGKDLME 326
 QY 329 VARRFNLPGV--RQGYGLTETSAILITPE--GDOPKASGKVPVLPFKKVIDLDTKTL 384
 Db 327 CGR--NIPNVLMOGGMETGCIIVSEVDRPLGRKSSGAGMLAPVEVLAQIYSEVETKSG 384
 QY 385 GPNRGEVCKGPMKKGVDNDEATRETIIDEGWLTHTDIDIGYDEKHFVYDRLKSLI 444
 Db 385 PPNQOGEIWWRGPMKKGVLNPNQAFKETIDKSKSWHTGDLGTFNEDGMLVYVDRIKELI 444
 QY 445 KYKGVOVPAELESVLLQHNIFEDAGVAGVDPPIAGCLPGAVVVLKKGKSMTEKEVMDV 504
 Db 445 KYKGVOVPAELESVLLQHNIFEDAGVAGVDPPIAGCLPGAVVVLKKGKSMTEKEVMDV 504
 QY 505 ASOVNAKRLRGVRFVDEVPKGLTGKIDKKAIRLEIKKIPVAM 548
 Db 505 AKQVAPYKRLR-RVSFLSLVPRKSAAKI---LRRELVOQVRSKM 544

RESULT 14

707908
 4-coumarate--CoA ligase (EC 6.2.1.12) 2 - western balsam poplar x cottonwood
 C:Species: Populus trichocarpa x Populus deltoides (western balsam poplar x cottonwood)
 C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 11-May-2000
 C:Accession: 707908
 R:Allina, S.M.; Pri-Hadash, A.; Theilmann, D.A.; Ellis, B.E.; Douglas, C.J.
 Plant Physiol. 116, 743-754, 1998
 A:Title: 4-coumarate:coenzyme A ligase in hybrid poplar. Properties of native enzymes
 A:Reference number: Z16208; MUID:98150279
 A:Accession: 707908
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-548 <ALU>
 A:Cross-references: EMBL:AF008183; NID:g2911796; PIDN:AC039365.1; PID:g2911797
 C:Genetics:
 A:gene: 4CL2
 C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
 C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis
 F:74-53/Domain: acetate--CoA ligase homology <ACLU>

Query Match 26.7%; Score 753.5; DB 2; Length 548;
 Best Local Similarity 36.0%; Pred. No. 8.6e-46;
 Matches 186; Conservative 86; Mismatches 193; Indels 51; Gaps 12;

QY 46 NALMGVDYAYEYLEKSCGLKNGVLDGRIALCSENCPEFFIYVLAGLFTGVGA 105
 Db 47 NGPTDITTYADVLTSSKVASGLKGLGQGDVILLILNOSPFEVAFIAGASTIGSS 106
 QY 106 PTNELYTLRELVSIGISKPTIVSSKKGLDKVITYQKIVTAIKTIYILDSKYDVRGYS 165
 Db 107 TANPPTSAELAKQATASKAKLIITHAAYAEKVOQFQENQHVH-ITMIDSLT----- 158
 QY 166 MDNFIRKATTPGKSSSKTYEVNRKEQVALIMNSGSGTLPKQVOLTREHIVYR----- 220
 Db 159 -ENCLHFSLELTSSDENELPTVKI--KPDIDMALPYSSGTTGLPKGVMLTHKGLVSAQOV 216
 QY 221 -----FSHADPIYGNQVSPGTAILLVPPFHGFGMFTTLGYLTCGFR---IWMUTK 269
 Db 217 DGENNLYFHERD-----VILCVLPFH---IYLSNPFVLCGLRAGSAILVMOK 262
 QY 270 FDETFELTKLDYKCSSVILVPTLFAILNRSFLDKYDLSNIVETASGAPLSKETGEAV 329
 Db 263 EDTVALMDLVQKYKVTIAPLVPLICLAIKSPVQDYDLSIRIVLISGAPLGELEDTV 322
 QY 330 ARFNLPGVR--QGYGLTETSAILI-----TPSDDKPGASGKVPVLPFKKVIDLDTK 382
 Db 323 --RAFLPNAKLGQGGMTAEAGPVIAMCLAFKPEPEIKSGAGCTVYRNAEKIYDPETGE 380
 QY 383 TLGPNRGEVCKGPMKKGVDNDEATRETIIDEGWLTHTDIDIGYDEKHFVYDRLKS 442
 Db 381 SOPRKRKTEGICRGQIMKGYLNDEPEATERITIDKGMWLTHTDIDIGYIDDE-LFIYDRKE 439
 QY 443 LKRYGYVPAELESVLLQHNIFEDAGVAGVDPPIAGCLPGAVVVLKKGKSMTEKEVMD 502
 Db 440 LKRYGYVPAELESVLLQHNIFEDAGVAGVDPPIAGCLPGAVVVLKKGKSMTEKEVMD 499
 QY 503 YVASOVNAKRLRGVRFVDEVPKGLTGKIDKKAIR 538
 Db 500 YISKOVIFYKRI-GRVFTTEALPKAPSGKILKDKLR 534

RESULT 15

703789
 4-coumarate--CoA ligase (EC 6.2.1.12) 4CL2 - common tobacco
 N:Alternative names: 4-coumaroyl-CoA synthetase
 C:Species: Nicotiana glauca (common tobacco)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
 C:Accession: 703789
 R:Lee, D.; Douglas, C.J.
 Plant Physiol. 112, 193-205, 1996
 A:Title: Two divergent members of a tobacco 4-coumarate:coenzyme A ligase (4CL) gene

A;Reference number: Z15086; MUTID:96416441

A;Accession: T03789

A; Status: preliminary; translated from GB/EMBL/DBJ

A; molecule type: mRNA

A;Residues: 1-542 <LEE>

A/Cross-references: EMBL:U50846; NID:g1663723; PIDN:AAB18638.1; PID:g1663724

C; Genetics:

A; Gene: 4CL2

C;Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
C;Keywords: acetyl-CoA ligase; coumarate 4-flavonoid biosynthesis

Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis

F;73-532/Domain: acetate--CoA ligase homology <ACL>

Query Match	26.48;	Score 746.5;	DB 2;	Length 542;
-------------	--------	--------------	-------	-------------

Best Local Similarity 34.98; Pred. No. 2.7e-45;

Matches 189; Conservative 90; Mismatches 200; Indels 63; Gaps 12;

```

QY 32 YMDRYAKGAJAFN-----ALGVD---VTYAELEKSSCGCEALKNKGLVYDGR 79
Db 20 YTPMHLPHLHSCFENISEFSSRPLINGANKQIITYADVELNSKRYAAAGLHKQIQPKDT 79
QY 80 IALCENCEEFFIYVLAGLFIQGVAPATNELIYTLRELVHSLGSLKSPYIVSSKGLDKVI 139
Db 80 IMILLPNSEFVAFIGASYLGAISTMANPLETFAEVVQAKAKSSAKTIITYQACHVKNV- 138
QY 140 TVQKTVTAIKTIVLDSKVDRKYQSDMNFK---KNTPOGEKGSFKT-----V 186
Db 139 -----KDY---AFENDYKILCIDSAPEGCLHFSVLTQANEHDIPEV 176
QY 187 EVNRKEOYALIMNSSGNGLPKQVULHENHYRFRSHARPIDYN-QVSGTALITYVPF 245
Db 177 EIQPDVDVAL-PYSSGTTGLPKGMALHKGGLVTSVAQOVGENPNLYIHSEDWLCVLP 235
QY 246 HHGGMFTTLYLTCGR---IYMLTKFDEETPLKTLQDYKCSSYLVPTLFAILNRSE 301
Db 236 FH---IYSINSVLLCGLRVGAALIMQKFDVSLTLELIQKRYKTIIGFVFPYIATAKSP 292
QY 302 LLDKYDLSNLVEIASGAPLSKETIGEAVARFNLPGVROGYLGETTSAIIT---TPE 356
Db 293 MWDYDIDLSVRYTWSGAPLCKELETDYTRAKFPRAKLGCGMTEAGPVLAMCLAFAPK 352
QY 357 GDDKRGASGVVPLFKRKVIDLQKTKLGPBRREYCVKPMIMKGVYNDPEATREIIDE 416
Db 353 FEIKSGACGYVRAAEKHYIYDPTGNSLPRNOSCIDTRSDQIMKGLNPDEARTARTIK 412
QY 417 EGMHLTGDIQYDEKHEFFIYDRLSLKTIKGYOVPAALESYVLQHPNIFDAGVAGVPD 476
Db 413 EGMLYTGIGIYIDDDDELFIYDRLKELIKTKKGFQVAALEALLNHPNISDAAVVPMKD 472
QY 477 PIAGELPGAVVYLKKGSMTEKEVMDYVASOVSAKTLRGGVRFVDEVPKGLTGKIDGKA 536
Db 473 EOMAGEVPAPFVVRNSGSTIYEDVKDSTISQVLIYFFKRIK-RVFEVDVAIPKSPSGKILKD 531
QY 537 IR 538
Db 532 LR 533

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:53:15 ; Search time 64.96 Seconds
(without alignments)
1233.948 Million cell updates/sec

Title: US-09-581-241-4

Percent score: 2823
Sequence: 1 MEMMENDENIYGPPEPTPL.....TKIDGKAIKRLKKPVAKM 548

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2793	98.9	548	5	027321 luciola lat
2	2792	98.9	548	5	027348 luciola lat
3	2343.5	83.0	548	5	025118 hotaria par
4	2335.5	82.7	548	5	026304 luciola min
5	1965.5	69.6	547	5	027688 lampyris no
6	1941.5	68.8	550	5	027755 photinus py
7	1918.5	68.0	550	5	027758 photinus py
8	1898	67.2	548	5	026076 pyrococella
9	1890	67.0	548	5	026076 pyrococella
10	1831	64.9	552	5	094697 photuris pe
11	1825	64.6	552	5	094696 photuris pe
12	1623.5	57.5	545	5	027757 photuris pe
13	1459.5	51.7	545	5	094408 phrixothrix
14	1330	47.1	546	5	094408 phrixothrix
15	1025	36.3	544	5	094408 phrixothrix
16	775.5	27.5	546	10	094408 phrixothrix
17	773	27.4	636	10	042879 lithospermu
18	772	27.3	562	10	094408 phrixothrix
19	758	26.9	557	10	048869 populus dal

20	755.5	26.8	548	10	048868	048868 populus dal
21	755	26.7	544	10	09M0X9	09M0X9 arabidopsis
22	748.5	26.5	580	10	09LMV8	09LMV8 arabidopsis
23	748	26.5	544	10	09LMV8	09LMV8 rubius idae
24	744	26.4	543	10	09LMV8	09LMV8 rubius idae
25	744	26.4	585	10	042880	042880 lithospermu
26	740.5	26.2	542	10	042943	042943 nicotiana t
27	735.5	26.1	597	5	09Y158	09Y158 drosophila
28	734.5	26.0	535	10	081139	081139 populus tre
29	734.5	26.0	545	10	08SMW8	08SMW8 solanum tub
30	721	25.5	1549	10	09LMV7	09LMV7 arabidopsis
31	720	25.5	544	5	019339	019339 caenorhabdi
32	715	25.3	570	10	081140	081140 populus tre
33	712	25.2	522	2	09K3W1	09K3W1 streptomyce
34	693	24.5	556	10	09M7S2	09M7S2 lolium pere
35	684	24.2	570	10	09LMV6	09LMV6 arabidopsis
36	676.5	24.0	542	10	09LMV2	09LMV2 arabidopsis
37	675.5	23.9	570	10	09M7S3	09M7S3 lolium pere
38	674	23.9	591	10	09LMV1	09LMV1 rubius idae
39	666.5	23.6	542	10	09FOX7	09FOX7 capsicum an
40	657	23.3	557	10	09M7S1	09M7S1 lolium pere
41	649	23.0	544	10	09M7S1	09M7S1 arabidopsis
42	635	22.5	552	1	028762	028762 archaeoglob
43	634.5	22.5	566	10	049414	049414 arabidopsis
44	602	21.3	566	2	09K8A2	09K8A2 bacillus ba
45	599.5	21.2	569	1	028502	028502 archaeoglob

ALIGNMENTS

RESULT 1
ID 027321 PRELIMINARY; PRT; 548 AA.
AC 027321;
DT 01-NOV-1996 (TREMUREL. 01, Created)
DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)
DT 01-JUN-2001 (TREMUREL. 17, Last annotation update)
DE LUCIFERASE.
OS Luciola lateralis (Firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharidea; Lampyridae; Luciola.
OX NCBI_TaxID=7052;
RN [1]
RP SEQUENCE FROM N.A.
RA Cho K., Choi Y., Boo K.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-MOUJ; TISSUE=ABDOMEN;
RC Cho K.H., Choi Y., Boo K.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN-MOUJ; TISSUE=ABDOMEN;
RC Cho K., Choi Y., Boo K.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z69619; CAA93444.1; -;
DR EMBL: U49181; AAA91471.1; -;
DR EMBL: U51019; AAB00229.1; -;
DR HSSP: P08659; ILCT.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP-BINDING; 1.
SQ SEQUENCE 548 AA: 60048 MW: E26C6C26F423E5E8 CRC64;

Query Match 98.9%; Score 2793; DB 5; Length 548;
Best Local Similarity 98.7%; Pred. No. 5.6e-185;
Matches 541; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 MEMMENDENIYGPPEPTPIEGSGAGLKRKYMDRVAKGLAIAFTNLGTGYDYAYEYLE 60

- Query Match 98.9%; Score 2792; DB 5; Length 548;

Query Match	83.08; Score 2343.5; DB 5; Length 548;
-------------	--

Best Local Similarity 81.9%; Pred. No. 7e-154;
Matches 447; Conservative 47; Mismatches 51; Indels 1; Gaps 1;

```

OY 4 MENDENIYVGEPEFYPIEBGSAGAOQLRKYMDRYAKLGAIAFNALGVDYTYAEYLEKSC 63
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 MEKENVYVGPFPYPIEBGSAGIOQLKYMHOYAKLGAIAFNALGVDIYVOEYFDITC 62
OY 64 CLEALKNYGLVYDGRIALCSENCEEFFIPVLAGLIGVAVAPNTEIYTLRELHNSLGIS 123
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 63 RLAEAMKNGKMGKOGSTIALCSENCEEFFIPVLAGLIGVAVAPNTEIYTLRELHNSLGIA 122
OY 124 KPTIVSSKKGDLVITVOKTVAITKIYIILDSKYDYRGYOSDMNFIRKNTPOGFGSSSF 183
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 123 OPTIVSSRKGLPKVLEQKTVCIKTIYILDSKVNFGHDCMETFIRKHEVLEGPPTSF 182
OY 184 KTYEV-NRKEQVALLMNSSGSTGLPKGVQLTNENLVTRSHARDPIYGNQVSPGTAILTV 242
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 183 VPLDVKNRKHQVALLMNSSGSTGLPKGVRIITHGAVTRFSHAKDPITYGNQVSPGTAILTV 242
OY 243 VPFHNGFMFTTIGYLTGCFRIYMLTKFDEEFLKTLQDYKCSSVILVPTLFALINRSEL 302
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 243 VPFHNGFMFTTIGYLTGCFRIYMLTKFDEEFLKTLQDYKCSSVILVPTLFALINRSEL 302
OY 303 LDKYDLSNLVEIASGAPLAKSKEIGEAVARFNLPGVROGYGLTETTSAILTPEGDDKPG 362
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 303 LDKYDLSNLVEIASGAPLAKSKEIGEAVARFNLPGVROGYGLTETTSAILTPEGDDKPG 362
OY 363 ASGVVPLFKAKYIDLDTKTKTLGPNRGEVCYKGPMLKGYVNDPNEATREIIDEBSMLHT 422
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 363 ASGVVPLFKAKYIDLDTKTKTLGPNRGEVCYKGPMLKGYVNDPNEATREIIDEBSMLHT 422
OY 423 GDIGYDEEKHFPIYDLRLKSLIKYGYOVPAELESVLLQHPNIFPDAGVAGVDPPIAGEL 482
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 423 GDIGYDEEKHFPIYDLRLKSLIKYGYOVPAELESVLLQHPNIFPDAGVAGVDPPIAGEL 482
OY 483 PGAVVVLKKGSMTEKEVMDYVASOVSNARKLRGCVRFVDEVPKGLTGKIDKAIKREILK 542
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 483 PGAVVVLKKGSMTEKEVMDYVASOVSNARKLRGCVRFVDEVPKGLTGKIDKAIKREILK 542
OY 543 KPVAKM 548
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 543 KPVAKM 548
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 4 PRELIMINARY: PRT: 548 AA.

```

ID Q26304
AC Q26304;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE LUCIFERASE.
OS Luciola mingrelica (Southern Russian firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharoidea; Lampyridae; Luciola.
OX NCBI_TaxID=27446;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93277944; PubMed=8504162;
RA Devine J.H., Kutuzova G.D., Green V.A., Ugarova N.N., Baldwin T.O.;
RT "Luciferase from the east European firefly Luciola mingrelica: cloning
RT and nucleotide sequence of the cDNA, overexpression in Escherichia
RT coli and purification of the enzyme."
RL Biochim. Biophys. Acta 1173:121-132(1993).
DR EMBL: S61961; AAB26932.1; -.
DR HSSP: P08659; 1BA3.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP-BINDING; UNKNOWN.1.
SEQUENCE 548 AA: 60494 MW: 62089533BOFFBC423 CRC64:

```

Query Match 82.7%; Score 2335.5; DB 5; Length 548;
Best Local Similarity 81.7%; Pred. No. 2.5e-153;
Matches 446; Conservative 47; Mismatches 52; Indels 1; Gaps 1;

```

OY 4 MENDENIYVGEPEFYPIEBGSAGAOQLRKYMDRYAKLGAIAFNALGVDYTYAEYLEKSC 63
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 MEKENVYVGPFPYPIEBGSAGIOQLKYMHOYAKLGAIAFNALGVDIYVOEYFDITC 62
OY 64 CLEALKNYGLVYDGRIALCSENCEEFFIPVLAGLIGVAVAPNTEIYTLRELHNSLGIS 123
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 63 RLAEAMKNGKMGKOGSTIALCSENCEEFFIPVLAGLIGVAVAPNTEIYTLRELHNSLGIA 122
OY 124 KPTIVSSKKGDLVITVOKTVAITKIYIILDSKYDYRGYOSDMNFIRKNTPOGFGSSSF 183
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 123 OPTIVSSRKGLPKVLEQKTVCIKTIYILDSKVNFGHDCMETFIRKHEVLEGPPTSF 182
OY 184 KTYEV-NRKEQVALLMNSSGSTGLPKGVQLTNENLVTRSHARDPIYGNQVSPGTAILTV 242
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 183 VPLDVKNRKHQVALLMNSSGSTGLPKGVRIITHGAVTRFSHAKDPITYGNQVSPGTAILTV 242
OY 243 VPFHNGFMFTTIGYLTGCFRIYMLTKFDEEFLKTLQDYKCSSVILVPTLFALINRSEL 302
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 243 VPFHNGFMFTTIGYLTGCFRIYMLTKFDEEFLKTLQDYKCSSVILVPTLFALINRSEL 302
OY 303 LDKYDLSNLVEIASGAPLAKSKEIGEAVARFNLPGVROGYGLTETTSAILTPEGDDKPG 362
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 303 LDKYDLSNLVEIASGAPLAKSKEIGEAVARFNLPGVROGYGLTETTSAILTPEGDDKPG 362
OY 363 ASGVVPLFKAKYIDLDTKTKTLGPNRGEVCYKGPMLKGYVNDPNEATREIIDEBSMLHT 422
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 363 ASGVVPLFKAKYIDLDTKTKTLGPNRGEVCYKGPMLKGYVNDPNEATREIIDEBSMLHT 422
OY 423 GDIGYDEEKHFPIYDLRLKSLIKYGYOVPAELESVLLQHPNIFPDAGVAGVDPPIAGEL 482
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 423 GDIGYDEEKHFPIYDLRLKSLIKYGYOVPAELESVLLQHPNIFPDAGVAGVDPPIAGEL 482
OY 483 PGAVVVLKKGSMTEKEVMDYVASOVSNARKLRGCVRFVDEVPKGLTGKIDKAIKREILK 542
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 483 PGAVVVLKKGSMTEKEVMDYVASOVSNARKLRGCVRFVDEVPKGLTGKIDKAIKREILK 542
OY 543 KPVAKM 548
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 543 KPVAKM 548
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 5

Q27688 PRELIMINARY: PRT: 547 AA.

```

ID Q27688
AC Q27688;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PHOTINUS-LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (PHOTINUS-LUCIFERIN
DE 4-MONOOXYGENASE (ATP-HYDROLYSING)) (FIREFLY LUCIFERASE) (LUCIFERASE).
OS Lampyris noctiluca.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharoidea; Lampyridae; Lampyris.
OX NCBI_TaxID=41311;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-LIGHT ORGAN;
RC MEDLINE=96190714; PubMed=8611152;
RA Sala-Newby G.B., Thomson C.M., Campbell A.K.;
RT "Sequence and biochemical similarities between the luciferases of the
RT glow-worm Lampyris noctiluca and the firefly Photinus pyralis."
RL Biochem. J. 313:761-767(1996).
CC -i- CATALYTIC ACTIVITY: PHOTINUS LUCIFERIN + O(2) + ATP - OXIDIZED
CC PHOTINUS LUCIFERIN + CO(2) + H(2)O + AMP + DIPHOSPHATE + LIGHT.
DR EMBL: X89479; CA61668.1; -.
DR HSSP: P08659; 1BA3.
DR InterPro: IPR000873; AMP-bind.
DR InterPro: IPR000215; Serpin.

```

DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
 KW Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 547 AA; 60365 MW; 8CB5653E0780D3EB CRC64;

Query Match 69.6%; Score 1965.5; DB 5; Length 547;
 Best Local Similarity 69.0%; Pred. No. 1e-127;
 Matches 372; Conservative 76; Mismatches 90; Indels 1; Gaps 1;

QY 4 MENDENIVYPERPEYPIEESGAGQLRKYMDRYAKL-GAIAFTNALTGVDYTAETLEKS 62
 DB 1 MEDAKNIMKGGPAPPYPLEDDTAGEQQLHAKMKRYAQAQVPGTIAFTDAHAEVITAEYSEMA 60
 QY 63 CCGEALKNKGLVYDGAIALCSCENCEEFIPVLAGLFIGVAVPTNITLRELVHSLGI 122
 DB 61 CRLAEIMKRRGGLQHHHIANVCSENSLOFMPVCGALFIGVASTNDIYMERELVHSLGI 120
 QY 123 SKPTIVSSKKGLDKYITVOKTVAIKTIVILDSKYDVRGYSMDNFIKKNTPOGFGSS 182
 DB 121 SPTIVSVSKRALOKILIGVOKKLPITOKIYILDSREDYMGKOSWYSPFESHLPAGENEYD 180
 QY 183 FKTEVNRKQVALINMSGSTGLPKGVOLTHENLYTRFSHARDPIYGNVSPGTALITV 242
 DB 181 YIPDSFDRKATLALINMSGSTGLPKGVOLTHENLYTRFSHARDPIYGNVSPGTALITV 240
 QY 243 VPRHNGGMMTTLGYLTCGRYIMLTKFDEETFLKTLQDYKSSVILVPLFALINSEL 302
 DB 241 IPRHNGGMMTTLGYLTCGRYIMLTKFDEETFLKTLQDYKSSVILVPLFSPFAKSTL 300
 QY 303 LDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDDKPG 362
 DB 301 VDKYDLSNLHEIASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDDKPG 360
 QY 363 ASKVYVPLFAKVIDDTKTKTLCPNRRGEYCVKPMIMKGYVNPATREIIDEEGWLHT 422
 DB 361 ASKVYVPLFAKVIDDTKTKTLCPNRRGEYCVKPMIMKGYVNPATREIIDEEGWLHT 420
 QY 423 GDIGYDEEKHEFTVRLKSLIKYKGYOVPAPLESVLLQHPNIPFAGVAGVDPDTAGEL 482
 DB 421 GDIAVYDEKHFEFTVRLKSLIKYKGYOVPAPLESVLLQHPNIPFAGVAGVDPDTAGEL 480
 QY 483 PGAVVVLKKGKSTKEKVDVYASQVSNKRLRGVRFVDEVPKGLTGKIDGKAIREIL 541
 DB 481 PAAVVVLKKGKSTKEKVDVYASQVSNKRLRGVRFVDEVPKGLTGKIDGKAIREIL 539

RESULT 6
 ID 027755 PRELIMINARY; PRT; 550 AA.

DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE FIREFLY LUCIFERASE (EC 1.13.12.7) (PHOTINUS-LUCIFERIN 4-MONOOXYGENASE
 (AMP-HYDROLYSING)).
 GN LUC.
 OS Photinus pyralis (North American firefly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidea; Lampyridae; Photinus.
 OX NCBI_TaxID=7054;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AUTOGRAPHIA CALIFORNICA NUCLEAR POLYHEDROSIS VIRUS-SSE STRAIN
 RA Croizier G.;
 RL Submitted (FEB-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: PHOTINUS LUCIFERIN + O(2) + ATP -> OXIDIZED
 CC PHOTINUS LUCIFERIN + CO(2) + H(2)O + AMP + DIPHOSPHATE + LIGHT.
 DR EMBL; X84846; CA59281.1; -;
 DR HSSP; P08659; 1BA3.

DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 550 AA; 60731 MW; 0A1CC749D0C96ADD CRC64;

Query Match 68.8%; Score 1941.5; DB 5; Length 550;
 Best Local Similarity 68.0%; Pred. No. 4.0e-126;
 Matches 368; Conservative 75; Mismatches 97; Indels 1; Gaps 1;

QY 4 MENDENIVYPERPEYPIEESGAGQLRKYMDRYAKL-GAIAFTNALTGVDYTAETLEKS 62
 DB 1 MEDAKNIMKGGPAPPYPLEDDTAGEQQLHAKMKRYAQAQVPGTIAFTDAHAEVITAEYSEMA 60
 QY 63 CCGEALKNKGLVYDGAIALCSCENCEEFIPVLAGLFIGVAVPTNITLRELVHSLGI 122
 DB 61 VRLAEAMKRRYGLNHNHIANVCSENSLOFMPVCGALFIGVAVPTNITLRELVHSLGI 120
 QY 123 SKPTIVSSKKGLDKYITVOKTVAIKTIVILDSKYDVRGYSMDNFIKKNTPOGFGSS 182
 DB 121 SPTIVSVSKRALOKILIGVOKKLPITOKIYILDSKIDYQSFQSYTFVISHLPAGENEYD 180
 QY 183 FKTEVNRKQVALINMSGSTGLPKGVOLTHENLYTRFSHARDPIYGNVSPGTALITV 242
 DB 181 FVPESEFDRKATLALINMSGSTGLPKGVOLTHENLYTRFSHARDPIYGNVSPGTALITV 240
 QY 243 VPRHNGGMMTTLGYLTCGRYIMLTKFDEETFLKTLQDYKSSVILVPLFALINSEL 302
 DB 241 VPRHNGGMMTTLGYLTCGRYIMLTKFDEETFLKTLQDYKSSVILVPLFSPFAKSTL 300
 QY 303 LDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDDKPG 362
 DB 301 IDKYDLSNLHEIASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDDKPG 360
 QY 363 ASKVYVPLFAKVIDDTKTKTLCPNRRGEYCVKPMIMKGYVNPATREIIDEEGWLHT 422
 DB 361 ASKVYVPLFAKVIDDTKTKTLCPNRRGEYCVKPMIMKGYVNPATREIIDEEGWLHT 420
 QY 423 GDIGYDEEKHEFTVRLKSLIKYKGYOVPAPLESVLLQHPNIPFAGVAGVDPDTAGEL 482
 DB 421 GDIAVYDEKHFEFTVRLKSLIKYKGYOVPAPLESVLLQHPNIPFAGVAGVDPDTAGEL 480
 QY 483 PGAVVVLKKGKSTKEKVDVYASQVSNKRLRGVRFVDEVPKGLTGKIDGKAIREIL 542
 DB 481 PAAVVVLKKGKSTKEKVDVYASQVSNKRLRGVRFVDEVPKGLTGKIDGKAIREIL 540
 QY 543 K 543
 DB 541 K 541

RESULT 7
 ID 027758 PRELIMINARY; PRT; 550 AA.

DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE FIREFLY LUCIFERASE (EC 1.13.12.7) (PHOTINUS-LUCIFERIN 4-MONOOXYGENASE
 (AMP-HYDROLYSING)).
 GN LUC.
 OS Photinus pyralis (North American firefly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidea; Lampyridae; Photinus.
 OX NCBI_TaxID=7054;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AUTOGRAPHIA CALIFORNICA NUCLEAR POLYHEDROSIS VIRUS-SSE STRAIN
 RA Croizier G.;
 RL Submitted (FEB-1995) to the EMBL/Genbank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: PHOTINUS LUCIFERIN + O(2) + AMP - OXIDIZED
CC PHOTINUS LUCIFERIN + CO(2) + H(2)O + AMP + DIPHOSPHATE + LIGHT
DR EMBL: X84847; CAA59282.1; -
DR HSSP: P08659; 1BA3
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding, 1.
DR PROSITE, PS00455; AMP_BINDING; 1.
KW Oxidoreductase.
KW SEQUENCE 350 AA; 60602 MW; D963B300D030F119 CRC64;

Query Match	68.0%	Score 1918.5	DB 5	Length 550
Best Local Similarity	67.5%	Pred. No. 1.8e-124		
Matches 365	Conservative 76	Mismatches 99	Indels 1	Gaps 1

Oy	4	MENNDNIYYGGEPPYPIEESGAGQOLKRYMDRAKL-GALAFNNALTDVDTYAEFLKS	62
Oy	1	MEMDANKIKKGAPRPPLEDDTAGOQLHAKMKRRALVPGTIAFDALHENVITYAEFLKS	60
Oy	63	CCGLGALKNGLVYDGRITALCSENCEFFPIPLVLAGLFGYGVAPFMEITTLRELHNSLGI	122
Db	61	VRLEAMKRYGLNNHRIIVYCSSELOFFEMVYLGALFIGAVALPADIYNERELNSMI	120
Oy	123	SKPTVESSKGLDVLTVOKTVAIKTYILYLDSDYDRGQSDMDFIKKNTPOGFKGSS	182
Db	121	SQPIYVAFSKGLQKILNVQKKLPIIOKIIIMOSKTDYQGFOSGMFYFVSHLPGRFNE	180
Oy	183	KTYEVNKRKEDVALIMNSSGSTGLPKGVOLTHENLVTRSHADPIYGNVSGTALTY	242
Db	181	FVPESFDKDTIALIMNSSGSTGSPKVALPHRTACVRFSHADPIFGNOIIPDTALSY	240
Oy	243	VPHNFGFMPTTLYLNCGFPIVWLTFGDETFPLTKLDYDKCSSVILPPLFLALNRSE	302
Db	241	VPHNFGFMPTTLYLNCGFPIVWLTFGDETFPLTKLDYDKCSSVILPPLFLALNRSE	300
Oy	303	LDKYDLSNLVEIASGAPSLKEIGEAAVARRPMLPGVRQGYGLTEETSAIITPEGDDKRG	362
Db	301	IDKYDLSNLVEIASGAPSLKEIGEAAVARRPMLPGVRQGYGLTEETSAIITPEGDDKRG	360
Oy	363	ASGVKVPLEFKKAVLDLQTKTGLGNRRGGEVCGSPMLKGYVONPRTREITIDEEGMLH	422
Db	361	ANGKAVPEFEKAVVLDLQTKTGLGNRRGGEVCGSPMLKGYVONPRTREITIDEEGMLH	420
Oy	423	GDIGYVDEKFFHYDRLKSLDIKKGYQVAPPAELSVLLOHPNIFDAGVAPDPTAGEL	482
Db	421	GDIAWDEDEHFIYDRLKSLDIKKGYQVAPPAELSVLLOHPNIFDAGVAGLGDGAGEL	480
Oy	483	PGAVVVLKKGSMKEKEMDVAAQVSNARKLRGCVFVDEVPKGLTKIDGKAIREIIL	542
Db	481	PAVAVVVLKKGSMKEKEMDVAAQVSNARKLRGCVFVDEVPKGLTKIDGKAIREIIL	540
Oy	543	K 543	
Db	541	K 541	
RESULT 8			
ID	Q9GFP9	PRELIMINARY;	PRT; 548 AA.
AC	Q9GFP9;		
DT	01-MAR-2001	(TREMBLrel. 16, Created)	
DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)	
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)	
DE	LUCIFERASE.		
OS	Pyrococella rufo.		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;		
OC	Elateriformia; Cantharoidea; Lampyridae; Pyrococella.		
OX	NCBI_Taxid=71223;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	TISSUE=LIGHT ORGAN;		
RA	Lee K.S., Park H.-J., Bae J.S., Lee K.S., Shon H.D., Jin B.R.;		

RT "Molecular Cloning and Expression of cDNA encoding Inoliferase from the
 R1 firefly, *Pyrocoelia rufa*."
 R2
 R3
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF328553; AAC45439.1; -
 DR InterPro: IPR000873; AMP-bind.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00284; SERPIN; UNKNOWN 1.
 SQ
 SEQUENCE 548 AA, 60787 MW, FE6DDEB7DDDDFEA9B CRC64;

Query Match	67.2%;	Score 1898;	DB 5;	Length 548;
Best Local Similarity	68.0%;	Pred. No. 4.7e-123;		
Matches 367;	Conservative 72;	Mismatches 99;	Indels 2;	Gaps 2

OY	4	MEMO-ENTIVYPERPEYFIEEGSSGAGDIRKMYRATV-GAIFTALNGVUTYAETLEK	61
OY	1	MEMDSKIHMGHRSILMEDGTAGEOLHKMKRYAQPVTIAFTPAHAEVNITVSEYFEM	60
OY	62	SCCLGELAKNGLVVDRIALMSENCEEPFIRPLAGLFTGVAVPATNEITVLLRELVHSIG	121
Db	61	SCRLAETMKRRGELGLOHHIALVSENLSQDEMPRCALFTGVAVPATNTDINERELVNSLF	120
OY	122	ISKPTVYSSKKGLDKVITVQKTVTAIKTIVLIDSKVYDRGYQSDMNFIKKNTPDGFKGS	181
Db	121	ISQPTIVPSCSRALQKITLGVHKKPLPVYQKTVLIDSREDYMGKQMSYFIESHLPGFMEY	180
OY	182	SKFYEVNARKQOVALLMNSSSGTGLPRGVQLTENLVTFPSHARPIQNGVSPCTALT	241
Db	181	DYIPDSFRETATGALLMNSSSGTGLPRGVQLTIMANCVSRSHCRDVFEGNOIIPPTALT	240
OY	242	VVRFHNGGMMFTTGLYLTGCFRIYMLTKPDEEFYTLTDIDYCVSSYIIVPFLTALINSE	301
Db	241	VYRFHNGGMMFTTGLYLTGCFRIYMLRFEEDFLRSLQDYKQSLALVPLTFSEFAKST	300
OY	302	LIDKTDLSNLVEIYASGAPLISKELGEVAVNARFMLPCVROGCTGLTETSALITTPBGDKP	361
Db	301	LVDKTDLSNLHEIYASGAPLAKVEGEAVNARFKPLGIRQDGLTETSALITTPBGDKP	360
OY	362	GASGVVPELFAKAVITDITKTKTLGPNRNGEVCKGPMIMKGYDNPENARTEIIDEGMIL	421
Db	361	GACGKGVVPEFFAKIYDIDTGTCTGLGVNQRGLCYKGMIMKGYNNPEANALIDQGMIL	420
OY	422	TGSDIGYUDEEKHFPIVDRKLSLKYKGYVPPALESLVLOHPNIFDAGVAGVPPINGE	481
Db	421	SSDIIYVYRKDHFPIVDRKLSLKYKGYVPPALESLILOHFFIFDAGVAGIIPPDAGE	480
OY	482	LGGAVVVLKKGKSMTERKVMYVYASVSNMKNRKGGRFVDEVPKGLTJKIDGKAIREIL	541
Db	481	LPAAVVVLLEGKMMTEQDEVMYVAGVYATSKRLRGKVFYDEVPKGLTJKIDSKRIREIL	540
RESULT	9		
ID	026076	PRELIMINARY;	PRT; 548 AA.
AC	026076;		
DT	01-NOV-1996 (TREMBlurel_01, Created)		
DT	01-NOV-1996 (TREMBlurel_01, Last sequence update)		
DT	01-JUN-2001 (TREMBlurel_17, Last annotation update)		
DE	LUCIFERASE.		
OS	Pyrococella mlyako.		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyplaga;		
OC	Elateriformia; Cantharoidea; Lampyridae; Pyrococella.		
OX	NCBI_TaxID=39363;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=WHOLE BODY;		
RX	MEDLINE=96061635; PubMed=7480137;		
RA	Ohmura Y., Ohba N., Toh H., Tsuji F.I.;		
RT	"Cloning, expression and sequence analysis of cDNA for the luciferases		
RL	from the Japanese fireflies", Pyrococella mlyako and Holaria parvula.";		
	Photochem. Photobiol.		62:309-313(1995).

DR EMBL: L39928; AAC37254.1; -.
 DR HSSP: P08659; 11CT1.
 DR InterPro: IPR000873; AMP-bind.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 DR PROSITE: PS00284; SERPIN; UNKNOMN; 1.
 DR SEQUENCE 548 AA; 60956 MW; CD37F63E982A9401 CRC64;

Query Match 67.0%; Score 1890; DB 5; Length 548;
 Best Local Similarity 67.8%; Pred. No. 1,7e-122;
 Matches 366; Conservative 72; Mismatches 100; Indels 2; Gaps 2;

QY 4 MEND-ENIVYGEPEFYPIEESGAGALRKYMRYAKL-GAIAFTNALGVDTVAEYLEKS 61
 DB 1 MEDDSKIMHGHHSILMEDGTAGEOLHKMKRYAQPGLIATDHAENVNITYSEYFEM 60
 QY 62 SCLGALNKGVLVDGRALALSCENCEEFFIPVLAGLFGVAPNTNITRELVSIG 121
 DB 61 SCRLAETMKRYGIGLOHIIAVCSETSLQFPMYVCAFLIGVAPNTINDIYNERELVNSLF 120
 QY 122 ISKPTIVSSKGLDKVITVOKTVAIKTIVILDSKVDYRGVQSMDFIKKTPQGFKGS 181
 DB 121 ISOPTIVFCSKRALQKILGVQKLPVYQKIVILDSKEDYMGQSMYFTESHLPAGFNEY 180
 QY 182 SFKTEVNRKEQVALIMNSSGSTGLPKGVOLTHENLVTRFSHARDPIYGNVSPGTALITV 241
 DB 181 DYPDSEDRFATALLIMNSSGSTGLPKGVOLTHENLVTRFSHARDPIYGNVSPGTALITV 240
 QY 242 VPFHGFMTFTLLGYTCGFRIVMLTKDEDETEFLTKLQDYKCSSVILVPTLALINRSE 301
 DB 241 VPFHGFMTFTLLGYTCGFRIVMLTKDEDETEFLTKLQDYKCSSVILVPTLALINRSE 300
 QY 302 LDKYDLSNLVEIASGAPLSKEIGEAVARRENLPVGRQGVGLTETSALITTPBGDDKP 361
 DB 301 LVQKYDLSNLHEIASGAPLSKEIGEAVARRENLPVGRQGVGLTETSALITTPBGDDKP 360
 QY 362 GASGVVPLFKAVYIDLTKTKLGNRRGECVCKPMKKGVDNPEATREIIDEEGWLH 421
 DB 361 GAGGVVPLFKAVYIDLTKTKLGNRRGECVCKPMKKGVDNPEATREIIDEEGWLH 420
 QY 422 TGDIGYDEKHEFIVDRKSLIKYKGYVPAPELESVILQHPNIFDAGVAGVDPPIAGE 481
 DB 421 SGDIATYDDGHHFIVDRKSLIKYKGYVPAPELESVILQHPNIFDAGVAGVDPPIAGE 480
 QY 482 LPGAVALLKKSMTKEKEMDYVASQVSNARKLRGVRFEVDEPKGLTGKIDGAIREIL 541
 DB 481 LPGAVALLEKGMTEQEMDYVAGVYTSKRLRGVKEVDEPKGLTGKIDSKIREIL 540

RESULT 10
 Q94697

PRELIMINARY; PRT; 552 AA.

ID Q94697
 AC Q94697;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE LUCIFERASE.
 GN LUC.
 OS Photuris pennsylvanica.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharidae; Lampyridae; Photuris.
 OX NCBI_TaxID=41716;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LANTERN;
 RA Zeno S., Shiraiishi S., Inouye S., Saigo K.;
 RT "Cloning, nucleotide sequence and expression of two cDNAs encoding for
 luciferase from Photuris firefly.";
 RT Submitted (NOV-1993) to the EMBL/Genbank/DBJ databases.
 DR EMBL: D25416; BAA05006.1; -.

DR HSSP: P08659; 11CT1.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 DR SEQUENCE 552 AA; 60970 MW; CAE6179AE26561DE CRC64;

Query Match 64.9%; Score 1831; DB 5; Length 552;
 Best Local Similarity 62.7%; Pred. No. 2.1e-118;
 Matches 339; Conservative 91; Mismatches 109; Indels 2; Gaps 2;

QY 4 MENDENIVYGEPEFYPIEESGAGALRKYMRYAKL-GAIAFTNALGVDTVAEYLEKS 62
 DB 1 MEIENNILIGPPYPLDEGTAGEOLHRAISRYAARVETLATYVTHELVTEYFEFIDVT 60
 QY 63 CCLGALNKGVLVDGRALALSCENCEEFFIPVLAGLFGVAPNTNITRELVSIG 122
 DB 61 CRLAAMKNNGYIGLOHTISVCSENCVQFPMFICALYGVATAPNTINDIYNERELVNSLSI 120
 QY 123 SKPTIVSSKGLDKVITVOKTVAIKTIVILDSKVDYRGVQSMDFIKKTPQGFKGS 182
 DB 121 SOPTIVFTSRNSLQKILGVQSRPLPIKKIITLIDKRDYLGVQSMDFIKKTPQGFKGS 180
 QY 183 FKTVEVNRKEQVALIMNSSGSTGLPKGVOLTHENLVTRFSHARDPIYGNVSPGTALITV 242
 DB 181 FKPLSFD-LDRVACIMNSSGSTGLPKGVOLTHENLVTRFSHARDPIYGNVSPGTALITV 239
 QY 243 VPFHGFMTFTLLGYTCGFRIVMLTKDEDETEFLTKLQDYKCSSVILVPTLALINRSE 302
 DB 240 VPFHGFMTFTLLGYTCGFRIVMLTKDEDETEFLTKLQDYKCSSVILVPTLALINRSE 299
 QY 303 LDKYDLSNLVEIASGAPLSKEIGEAVARRENLPVGRQGVGLTETSALITTPBGDDKP 362
 DB 300 VDKYDLSNLHEIASGAPLSKEIGEAVARRENLPVGRQGVGLTETSALITTPBGDDKP 359
 QY 363 ASGVVPLFKAVYIDLTKTKLGNRRGECVCKPMKKGVDNPEATREIIDEEGWLH 422
 DB 360 AVGVVPLFKAVYIDLTKTKLGNRRGECVCKPMKKGVDNPEATREIIDEEGWLH 419
 QY 423 GDIGYDEKHEFIVDRKSLIKYKGYVPAPELESVILQHPNIFDAGVAGVDPPIAGE 482
 DB 420 GDIGYDEKHEFIVDRKSLIKYKGYVPAPELESVILQHPNIFDAGVAGVDPPIAGE 479
 QY 483 PGAVVPLKKSMTKEKEMDYVASQVSNARKLRGVRFEVDEPKGLTGKIDGAIREIL 542
 DB 480 PGAVVPLKKSMTKEKEMDYVASQVSNARKLRGVRFEVDEPKGLTGKIDGAIREIL 539
 QY 543 K 543
 DB 540 K 540

RESULT 11
 Q94696

PRELIMINARY; PRT; 552 AA.

ID Q94696
 AC Q94696;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE LUCIFERASE.
 GN LUC.
 OS Photuris pennsylvanica.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharidae; Lampyridae; Photuris.
 OX NCBI_TaxID=41716;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LANTERN;
 RA Zeno S., Shiraiishi S., Inouye S., Saigo K.;
 RT "Cloning, nucleotide sequence and expression of two cDNAs encoding for
 luciferase from Photuris firefly.";
 RT Submitted (NOV-1993) to the EMBL/Genbank/DBJ databases.

DR EMBL: D25415; BAA05005.1; -
 DR HSSP: P08659; 1BA3.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-BINDING; 1.
 DR SEQUENCE 552 AA; 61000 MW; 85C14ED52BD5366A CRC64;

Query Match 64.6%; Score 1825; DB 5; Length 552;
 Best Local Similarity 62.5%; Pred. No. 5.3e-118;
 Matches 338; Conservative 92; Mismatches 109; Indels 2; Gaps 2;

QY 4 MENDENIYVGEPEPPYPIEGSGAQLRKRYMDRYAKL-GALAFNALTGVDYTYAEYLEKS 62
 DB 1 MSTENNLIGPPYPLPEEGTAGELHRAITRYAAVPGTLATYDTHVELEYTKPEFLDVT 60
 QY 63 CCLGALKNYGLVVDGRALACSECEFFIPVLAGLIGVAVPNTNELLRELVHSLGI 122
 DB 61 CRLEAMKNGYGLQHTISVCSENCVOFEMPVCAALYIGATAPTNDIYNERELYNLSLI 120
 QY 123 SKPTIVSSKGLDKVITVOKTVAITKIYVILDSKVDYRGQSDMDFIKNTPOGFGSS 182
 DB 121 SQPIVVFTRSLKRIKIGVSRVYIKIIMLDTKKYLIGQSMQSKKEHVPANFNVSA 180
 QY 183 FKTYEVNKEQVALIMNMSGSTGLPKGVOLTHENLVTRFSHARDPIYGNQVSPGATILTY 242
 DB 181 FKPLSFD-LDRVACIMNMSGSTGLPKGVPISSHRTYTRFSDCRDPVPGNOIIPDTILCA 239
 QY 243 VPFHNGRMFTTLGVLGCFRIYMLTKFDEFTLKTLDYKCSSVILVPLFALLNSEL 302
 DB 240 VPFHNAFTFNLGYIICGFHVIMRYRNEHFLDTLDYKQSSALIVPVYLAELANPL 299
 QY 303 LDKYDLSLVLIASGAPLSEKEGEAVARFNLPGVROGGLTETTSALITPEGDDKPG 362
 DB 300 VDKTDLHLHLIASGAPLSEKEISEIAKRRKGLINOGIGLITTTCAIYTAGGEKPG 359
 QY 363 ASGRVPLFKAKVIDLDTKTKLGPNNRGEVCGPMLKGYVDNPEARTREIIDEGMLHT 422
 DB 360 AVGVAVPYSLKVLDTLNLGKTLGPNEGEICFTGPMIMKGYINNPENAREIIDEGMHTS 419
 QY 423 GDIGYDEKHFYIVDLKSLIKTKGYVPPAELESVLLQHPNIFDQVAVGPPINAGEL 482
 DB 420 GDIGFDDGHVYIVDLKSLIKTKGYVPPAELEALLQHPFIEDAGVAVGPPVAVGDL 479
 QY 483 PGAVVVLKKGSMTEKEVMDYVASOVSNARKLRGVRVDPVPGKLTGKIDKAIKIRELK 542
 DB 480 PGAVVVLKKGSMTEKEIQLDYVAGVITSSKRLRGVRFVKEVPGFTGKIDTRIKIELI 539
 QY 543 K 543
 DB 540 K 540

RESULT 12
 ID Q2757 PRELIMINARY; PRT; 545 AA.
 AC Q2757; 002653;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DE LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).
 OS Photuris pennsylvanica.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidea; Lampyridae; Photuris.
 NC NCB1;taxid=41716;
 RX MEDLINE-9307756; PubMed-9165098;
 RA Ye L., Buck L.M., Schaefer H.J., Leach F.R.;
 RT "Cloning and sequencing of a cDNA for firefly luciferase from Photuris pennsylvanica";

RL Biochim. Biophys. Acta 1339:39-52(1997).
 CC - FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 562 NM (BY
 CC SIMILARITY).
 CC - CATALYTIC ACTIVITY: PHOTINUS LUCIFERIN + O(2) + ATP = OXIDIZED
 CC PHOTINUS LUCIFERIN + CO(2) + H(2)O + AMP + DIPHOSPHATE + LIGHT.
 CC - COFACTOR: REDUCES A MAGNESIUM ION (BY SIMILARITY).
 CC - SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
 CC - SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
 DR EMBL: U31240; AAB60897.1; -
 DR HSSP: P08659; 1LCT.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-BINDING; UNKNOWN_1.
 KW Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
 KM Peroxisome.
 FT SITE 543 545 MICROBODY TARGETING SIGNAL (POTENTIAL).
 SQ SEQUENCE 545 AA; 60649 MW; F0FE4B828047C26E CRC64;

Query Match 57.5%; Score 1623.5; DB 5; Length 545;
 Best Local Similarity 56.5%; Pred. No. 4.6e-104;
 Matches 307; Conservative 97; Mismatches 138; Indels 1; Gaps 1;

QY 7 DENIYVGEPEPPYPIEGSGAQLRKRYMDRYAKL-GALAFNALTGVDYTYAEYLEKSCL 65
 DB 3 DKNLIVGEPEFHLADGTACGEQMFALSRADISGIALTNATKENVLYEEFLKSLR 62
 QY 66 GELAKNYGLVVDGRALACSECEFFIPVLAGLIGVAVPNTNELLRELVHSLGISRP 125
 DB 63 AESFKRYGLKQNDTIYAVCSENGLOFELPLIASLYLIIAIPVADSKIERELHSLGIYVP 122
 QY 126 TIYESSKGLDKVITVOKTVAITKIYVILDSKVDYRGQSDMDFIKNTPOGFGSSPFT 185
 DB 123 RIIFCSKNFTQKLVNYSKLYETIILIDNEDLGCTGCLNFIQSDINDLMDKKEFP 182
 QY 186 VEYNRKQVALIMNMSGSTGLPKGVOLTHENLVTRFSHARDPIYGNQVSPGATILTYVP 245
 DB 183 NSPNRDDQVALVMSGSTGLGVSKVLMTHKNIVARSHCKDPFGNAIPTTALITVIFP 242
 QY 246 HHGFGFTTIGYTCGFRIVMLTKFDEFTLKTLDYKCSSVILVPLFALLNRSGLDK 305
 DB 243 HHGFGMTTIGYTCGFRVAMLMHTFEKLELQSLQYKVESLTLVPTLMAFPKSLVYK 302
 QY 306 YDLSNLVEISGAPLSEKEGEAVARFNLPGVROGGLTETTSALITPEGDDKRGAG 365
 DB 303 YDLSHLKELIASGAPLSEKEGEVKKRFLNFRKGTGLTETTSALITPDDIVRGSTG 362
 QY 366 KVVPLFKAKVIDLDTKTKLGPNNRGEVCGPMLKGYVDNPEARTREIIDEGMLHTGDI 425
 DB 363 KIYVFAVKKVVDPTGTGILGPNETGELYFGDMIMKSYNNNEATKAIINKGWLRSDDI 422
 QY 426 GYVDEKHFYIVDLKSLIKTKGYVPPAELESVLLQHPNIFDQVAVGPPINAGELPRA 485
 DB 423 AYVNDGHFIYIVDLKSLIKTKGYVPPAELEIGILLQHPYIDAGVTGLIDEAAGELPRA 482
 QY 486 VVVLKKGSMTEKEVMDYVASOVSNARKLRGVRVDPVPGKLTGKIDKAIKIREIKRV 545
 DB 483 GVVVQTKYINQIVQNFVSSQVSTAKWLRGVKFLDEIPKSTGKIDKYLRLQMEFKIK 542
 QY 546 AKM 548
 DB 543 SKL 545

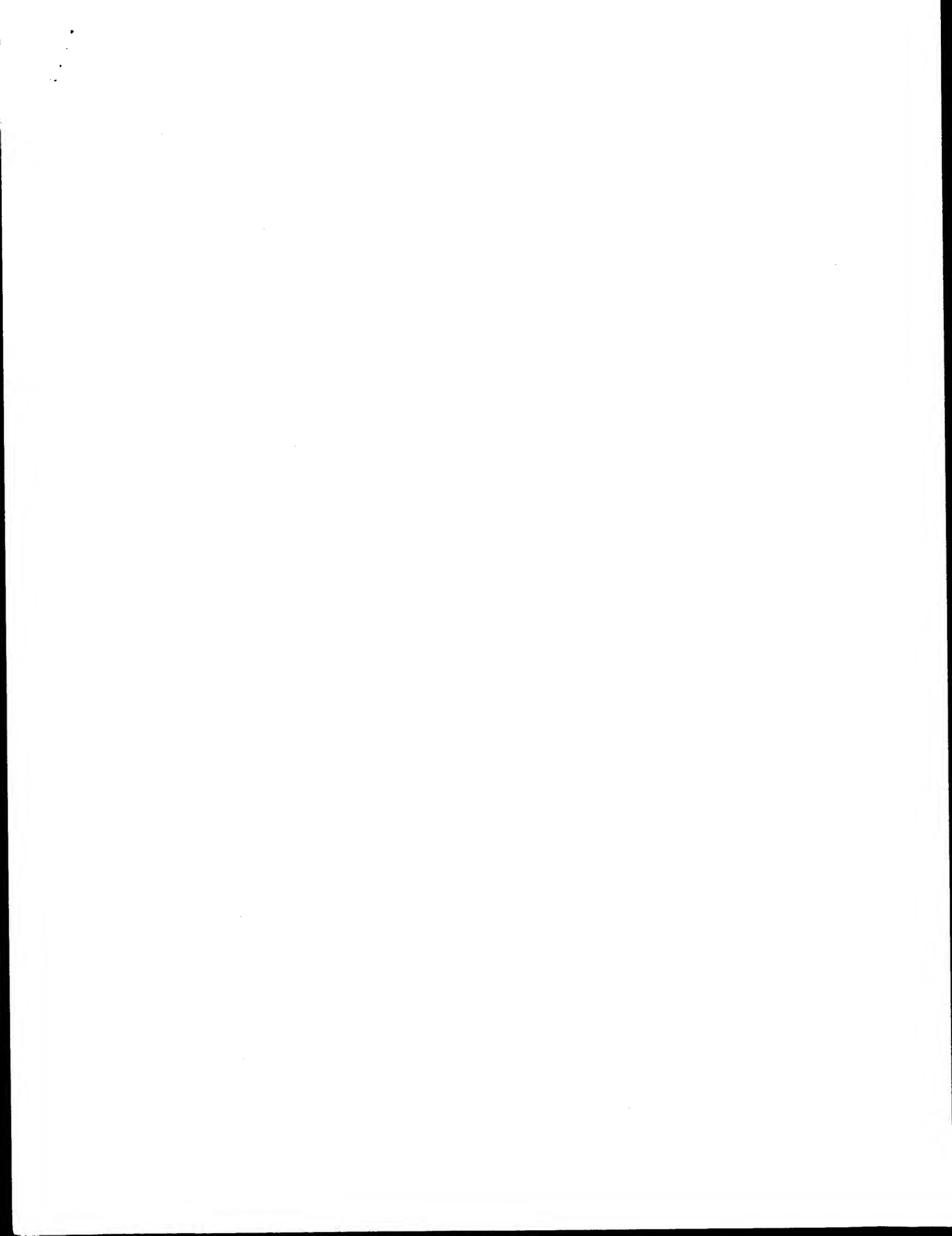
RESULT 13
 ID Q9U408 PRELIMINARY; PRT; 545 AA.
 AC Q9U408;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DE LUCIFERASE.
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)


```

Db      124 KPIIFASITITDRVAVAKSNKFKGITALSGSKRKNITDYLKELMEDEKFTQP--- 186
QY      179 KSSPKTEVANKKEQVALIMNMSGSTGLPKGVOLTHEMLYTRSHARDPIYGNQVSP-- 235
Db      181 ---DPTSPANKDEBVSILVSCSGTGLPKGVOLTMNLATLD-----SQIPTVI 229
QY      236 ---GTAITLVVPEHHGFMFTTGLVLTGCRFIMVLTKFDEFTPLTKDQKCSSVLVPT 292
Db      230 PMEVYTLITLVIPIWFAHFGCLTTLITACGAGALVYLLPRKEELFLSAIEKRVMMAFVPP 269
QY      293 LEAIIINRELLDKYDLSNLYEASGAPLSKEIEENAVARRNLPBGVNOGGLFTTSAIT 352
Db      290 LMVFLAKRPIYDKDLSLMLILCGAPLSKEDQIKERIGVFIKOGYGLSESTLSVL 349
QY      333 ITPBGDDKPGASGVVPLFKAVIDLTKTTLGNRGEVGVKPYMLKGVYVNPATRE 412
Db      350 VONNEFCRPGSVGLKVGIVAKVIDPDTGKLKGANERGEGLCFKDDGIMKGYITGDTSTQ 409
QY      413 IIDEGMLHTDIDIGYDEEKHFIVDRLSIKIKYGVYPAELESYLDHPNIFAGVA 472
Db      410 AI-KDGMHTSDIDIGYDDPEFFIVDRIKELIKKGIQVPAEELALLTNDXIKIDAAV 468
QY      473 GVPDPPIGELPGAVYVLKKGKSMTEKEVMDIVASQVSNARKLGGVAFVDEVPKGLTGKI 532
Db      469 GRPPEEGEELPLAVVAKQAVOLTENEVIOFVNDNAPSARKRLGGVIFVEIPIKNSGKI 528
QY      533 DGKAIRIELIKRPVAKM 548
Db      529 LRRILREMLKQKSKL 544

```

Query Match	36.3%	Score 1025;	DB 5;	Length 544;
Best Local Similarity	41.7%	Pred. No. 1.2e-62;		
Matches 232;	Conservative 97;	Mismatches 191;	Indels 36;	Gaps 11;
QY	NIYYGEPEYPIEBSA--GQLRKY-MDRAKAG-AIATPMTLGVDTYAEYLEKSCC	64		
	9	NIYYG----	PYTERQADSRSLGQYIIDKRYKSGCDRTVLVDAVNGVEYS-ASFMRKSTV	63
QY	65	-LGEALKNYGVADGRIALCSNCEEFPIPLAALFIVGVAAPNPNTYTLRELIVHSLGIS	123	
DB	64	RLATYIIQKIGVAKQNDVGVLSSENSVNFALAMFALAGATVAPLNTVYSPREVDHAINLS	123	
QY	124	KPIIVSSKKGDKVITVQKVTAKTIVILDS-----KDYKGYSDMNFIKKNTPGGF	178	



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:59:32 ; Search time 64.96 Seconds
(without alignments)
1233.948 Million cell updates/sec

Title: US-09-581-241-6

Perfect score: 2823
Sequence: 1 MNNMNDENIYGPPEPPPI.....TGKIDGKAIRKILKKPVAKM 548

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP.TREMBL.17:*
2: SP.Archea:*
3: SP.Bacteria:*
4: SP.Fungi:*
5: SP.Human:*
6: SP.Invertebrate:*
7: SP.Mhc:*
8: SP.Mammal:*
9: SP.Mhc:*
10: SP.Plant:*
11: SP.Rodent:*
12: SP.Virus:*
13: SP.Vertebrate:*
14: SP.Unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2793	98.9	548	5	Q27321 luciola lat
2	2792	98.9	548	5	Q27348 luciola lat
3	2343.5	83.0	548	5	Q25118 hotaria par
4	2335.5	82.7	548	5	Q26304 luciola min
5	1967.5	69.7	547	5	Q27688 lampyrin no
6	1941.5	68.8	550	5	Q27755 photinus py
7	1918.5	68.0	550	5	Q27758 photinus py
8	1900	67.3	548	5	Q9GPP9 pyrocoelia
9	1892	67.0	548	5	Q26076 pyrocoelia
10	1831	64.9	552	5	Q94697 photuris pe
11	1825	64.6	552	5	Q94696 photuris pe
12	1625.5	57.6	545	5	Q27757 photuris pe
13	1457.5	51.6	545	5	Q94048 phrixothrix
14	1332	47.2	546	5	Q94407 phrixothrix
15	1023	36.2	544	5	Q9VC66 atropisila
16	775.5	27.5	546	10	Q9C5H2 aradidopsis
17	771	27.3	636	10	Q42879 lithosperm
18	770	27.3	562	10	Q9FGW4 aradidopsis
19	756	26.8	557	10	Q48869 populus bal

20	754	26.7	544	10	Q9M0X9	Q9M0X9 aradidopsis
21	753.5	26.7	548	10	Q48868	Q48868 populus bal
22	748.5	26.5	580	10	Q9LMV8	Q9LMV8 aradidopsis
23	746	26.4	544	10	Q9L150	Q9L150 rubus idae
24	742	26.3	543	10	Q9LL49	Q9LL49 rubus idae
25	742	26.3	585	10	Q42880	Q42880 lithosperm
26	738.5	26.2	542	10	Q42943	Q42943 nicotiana t
27	737.5	26.1	597	5	Q9Y158	Q9Y158 drosophila
28	732.5	25.9	535	10	Q81139	Q81139 populus tre
29	732.5	25.9	545	10	Q9SMH8	Q9SMH8 solanum tub
30	720	25.5	544	5	Q19339	Q19339 caenorhabdi
31	719	25.5	1549	10	Q9LMV7	Q9LMV7 aradidopsis
32	714	25.3	522	2	Q9K3W1	Q9K3W1 streptomyce
33	713	25.3	570	10	Q81140	Q81140 populus tre
34	691	24.5	556	10	Q9M7S2	Q9M7S2 lolium pere
35	682	24.2	570	10	Q9LU36	Q9LU36 aradidopsis
36	674.5	23.9	542	10	Q9LQ12	Q9LQ12 aradidopsis
37	673.5	23.9	570	10	Q9M7S3	Q9M7S3 lolium pere
38	672	23.8	591	10	Q9LL51	Q9LL51 rubus idae
39	664.5	23.5	542	10	Q9FOY7	Q9FOY7 capsicum an
40	655	23.2	557	10	Q9M7S1	Q9M7S1 lolium pere
41	647	22.9	544	10	Q9FE44	Q9FE44 aradidopsis
42	635	22.5	552	1	Q28762	Q28762 archaeoglob
43	632.5	22.4	566	10	Q49414	Q49414 aradidopsis
44	600	21.3	566	2	Q9K8A2	Q9K8A2 bacillus ha
45	597.5	21.2	569	1	Q28502	Q28502 archaeoglob

ALIGNMENTS

RESULT 1
ID Q27321 PRELIMINARY: PRT: 548 AA.
AC Q27321;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE LUCIPERASE.
OS Luciola lateralis (Firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharoidea; Lampyridae; Luciola.
OX NCBI_TaxID=7052;
RN [1]
RP SEQUENCE FROM N.A.
RA Cho K., Choi Y., Boo K.;
RN Submitted (FEB-1996). to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cho K.H., Choi Y., Boo K.;
RN Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN-MUTU; TISSUE-ABDOMEN;
RA Cho K., Choi Y., Boo K.;
RN Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: Z69619; CA93444.1; -;
DR EMBL: U49181; AAA91471.1; -;
DR EMBL: U51019; AAB00229.1; -;
DR HSSP: P08659; ILCT.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP-BINDING; 1.
SQ SEQUENCE 548 AA: 60048 MW: E26CEC26FA23E5E8 CRC64;

Query Match 98.9%; Score 2793; DB 5; Length 548;
Best Local Similarity 98.7%; Pred. No. 1.2e-184;
Matches 541; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
OY 1 MNNMNDENIYGPPEPPPIEGSAGQALRKRYAKLGAIAFTALNGVDTTAYEIE 60

Best Local Similarity 81.9%; Pred. No. 1.4e-153;
Matches 447; Conservative 47; Mismatches 51; Indels 1; Gaps 1;

```

QY 4 MENDENIVGPEPPYPIEESGAGALRKYMRYAKGALAFNALGVDTYAEYLEKSC 63
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 3 MEKENVYVGLPYPIEESGAGIOLHKMQYAKGALAFNALGVDTYAEYLEKSC 62
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 64 CLGELAKNTGLVNDGRALALSCENCEEFPVLAGLFIGVAPNTNITLRELHSLGIS 123
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 63 RLAEAKNKGMMQEGTALSCENCEEFPVLAGLIGVAPNTNITLRELHSLGIS 122
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 124 KPTIVSSKKGLDKVITVQKTATTAITVILDSKVYRGYOSMDNFIRKNTPOGKSSSF 183
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 123 OPTIVSSRKGLPKYLEVQKTYCTITVILDSKVNFGHDCMETFIKKHVELGFPSTF 182
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 184 KTVETV-NRKEOVALIMSSSGTGLPKGYQLTHENITVRESHARDPIYGNVSPGTALITV 242
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 183 VPLDVKNRKHVALIMSSSGTGLPKGYRLTHGAVTRFESHAKDPIYGNVSPGTALITV 242
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 243 VPHHGFMTTGLTGLCGFRIVMLTKPDEEFTLKTLDYKCSSVILVPTLFAILNRSEL 302
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 243 VPHHGFMTTGLTGLCGFRIVMLTKPDEEFTLKTLDYKCTSVILVPTLFAILNRSEL 302
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 303 LDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGYGLTETTSATITPEGDDKPG 362
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 303 IDKFDLSNLTEIASGAPLAKGEVAVARFNLPGVROGYGLTETTSATITPEGDDKPG 362
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 363 ASGRVVPLEFKAVIDLTKTKTIGPNRGEVCYKGMIMKGYVNDPEATREIIDEQWMLHT 422
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 363 ASGRVVPLEFKAVIDLTKTKTIGPNRGEVCYKGMIMKGYVNDPEATREIIDEQWMLHT 422
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 423 GDIGYDEDEHFFIYDRLKSLIKYGYVPPAELESVLLQHPNIFDAGVAGVPPDAGEL 482
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 423 GDIGYDEDEHFFIYDRLKSLIKYGYVPPAELESVLLQHPNIFDAGVAGVPPDAGEL 482
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 483 PGAVVVLKKGKSMTEKEVMDYVASOVSNARKLGGVRFVDEVPKGLTGKIDGAKIREILK 542
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 483 PGAVVVLKKGKSMTEKEIYDVNSQVNNHRLRGVRFVDEVPKGLTGKIDAKVIREILK 542
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 543 KPQAKM 548
  || |||
Db 543 KPQAKM 548
  || |||

```

RESULT 4
ID Q26304 PRELIMINARY; PRT; 548 AA.
AC Q26304;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE LUCIFERASE.
OS Luciola mingrelia (Southern Russian firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriforma; Cantharoidea; Lampyridae; Luciola.
OX NCBI_TaxID=27446;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93277944; PubMed=8504162;
RA Devine J.H., Kutuzova G.D., Green V.A., Ugarova N.N., Baldwin T.O.;
RT "Luciferase from the east European firefly Luciola mingrelia: cloning
and nucleotide sequence of the cDNA, overexpression in Escherichia
coli and purification of the enzyme.";
RL Biochim. Biophys. Acta 1173:121-132(1993).
DR EMBL; S61961; AAB26932.1; -;
DR HSSP; P08659; 1BA3.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP BINDING; UNKNOWN_1.
DR SEQUENCE 548 AA; 60494 MW; 62C8953BDFB423 CRC64;

Query Match 82.7%; Score 2335.5; DB 5; Length 548;
Best Local Similarity 81.7%; Pred. No. 4.9e-153;
Matches 446; Conservative 47; Mismatches 52; Indels 1; Gaps 1;

```

QY 4 MENDENIVGPEPPYPIEESGAGALRKYMRYAKGALAFNALGVDTYAEYLEKSC 63
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 3 MEKENVYVGLPYPIEESGAGIOLHKMQYAKGALAFNALGVDTYAEYLEKSC 62
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 64 CLGELAKNTGLVNDGRALALSCENCEEFPVLAGLFIGVAPNTNITLRELHSLGIS 123
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 63 RLAEAKNKGMMQEGTALSCENCEEFPVLAGLIGVAPNTNITLRELHSLGIS 122
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 124 KPTIVSSKKGLDKVITVQKTATTAITVILDSKVYRGYOSMDNFIRKNTPOGKSSSF 183
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 123 OPTIVSSRKGLPKYLEVQKTYCTITVILDSKVNFGHDCMETFIKKHVELGFPSTF 182
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 184 KTVETV-NRKEOVALIMSSSGTGLPKGYQLTHENITVRESHARDPIYGNVSPGTALITV 242
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 183 VPLDVKNRKHVALIMSSSGTGLPKGYRLTHGAVTRFESHAKDPIYGNVSPGTALITV 242
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 243 VPHHGFMTTGLTGLCGFRIVMLTKPDEEFTLKTLDYKCSSVILVPTLFAILNRSEL 302
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 243 VPHHGFMTTGLTGLCGFRIVMLTKPDEEFTLKTLDYKCTSVILVPTLFAILNRSEL 302
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 303 LDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGYGLTETTSATITPEGDDKPG 362
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 303 IDKFDLSNLTEIASGAPLAKGEVAVARFNLPGVROGYGLTETTSATITPEGDDKPG 362
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 363 ASGRVVPLEFKAVIDLTKTKTIGPNRGEVCYKGMIMKGYVNDPEATREIIDEQWMLHT 422
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 363 ASGRVVPLEFKAVIDLTKTKTIGPNRGEVCYKGMIMKGYVNDPEATREIIDEQWMLHT 422
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 423 GDIGYDEDEHFFIYDRLKSLIKYGYVPPAELESVLLQHPNIFDAGVAGVPPDAGEL 482
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 423 GDIGYDEDEHFFIYDRLKSLIKYGYVPPAELESVLLQHPNIFDAGVAGVPPDAGEL 482
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 483 PGAVVVLKKGKSMTEKEVMDYVASOVSNARKLGGVRFVDEVPKGLTGKIDGAKIREILK 542
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 483 PGAVVVLKKGKSMTEKEIYDVNSQVNNHRLRGVRFVDEVPKGLTGKIDAKVIREILK 542
  ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 543 KPQAKM 548
  || |||
Db 543 KPQAKM 548
  || |||

```

RESULT 5
ID Q27688 PRELIMINARY; PRT; 547 AA.
AC Q27688;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE PHOTINUS-LUCIFERIN 4-MONOOXYGENASE (BC 1.13.12.7) (PHOTINUS-LUCIFERIN
4-MONOOXYGENASE (ATP-HYDROLYSING)) (PIRELY LUCIFERASE) (LUCIFERASE).
OS Lampyris noctiluca.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriforma; Cantharoidea; Lampyridae; Lampyris.
OX NCBI_TaxID=41311;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIGHT ORGAN;
RC MEDLINE=96190714; PubMed=8611152;
RA Sala-newby G.B., Thomson C.M., Campbell A.K.;
RT "Sequence and biochemical similarities between the luciferases of the
glow-worm Lampyris noctiluca and the firefly Photinus pyralis.";
RL Biochem. J. 313:761-767(1996).
DR EMBL; X89479; CAA61668.1; -;
DR HSSP; P08659; 1BA3.
DR InterPro; IPR000873; AMP-bind.
DR InterPro; IPR000215; Serpin.

CC -1- CATALYTIC ACTIVITY: PHOTINUS LUCIFERIN + O(2) + ATP = OXIDIZED
 CC PHOTINUS LUCIFERIN + CO(2) + H(2)O + AMP + DIPHOSPHATE + LIGHT.
 CC EMBL: X84847; CAA59282.1; -
 DR HSSP: P08659; 1BA3.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 DR Oxidoreductase.
 KM
 SQ SEQUENCE 550 AA; 60602 MW; D963B300D030F119 CRC64;

Query Match 68.0%; Score 1918.5; DB 5; Length 550;
 Best Local Similarity 67.5%; Pred. No. 3,1e-124;
 Matches 365; Conservative 76; Mismatches 99; Indels 1; Gaps 1;

QY 4 MENDENTVGPPEPPYPIEESGAGOLKRYMDRAKL-GAIAFTNALGVDTYAEYLEKS 62
 DB 1 MEDAKNIKGPAPFYPLEDGTAGEOLHKAKRYALVPGTIAFTDAHEVNTYAEYFEMS 60
 QY 63 CCLGEALKNYGLVVDGRIALCSENCEFFIPLVAGLFIGVAVPTNEITYLRELVSIG 122
 DB 61 VRLAEAKRKRIGLTNHRIYVCSNSLOFPMVCGALFIGVAVAPANDIYERELLSMNI 120
 QY 123 SKPTIVSSKKGLDKYTVOKTVAITVILDSKVDYRGYQSDNFIKNTPOGFKGSS 182
 DB 121 SQPTIVSVSKKGLQKILNVQKLPILQKILIMDSKTDYQGFQSMYFVTSHPGCFNEYD 180
 QY 183 FKTYEVNRKEQVALIMNSSGSTGLPKGVOLTHENITVTRSHADPIYGNVSPGTALTY 242
 DB 181 FVPESEFDRDITIALIMNSSGSTGPKGVAPLPHRTACVRFSHADPIFGNDIIPDTALISV 240
 QY 243 VPFHNGFMFTTGLYLCGFRIVMLTKFDEETFLKTDYDKCSSVILVPTLFAILNRSEL 302
 DB 241 VPFHNGFMFTTGLYLCGFRIVMLTKFDEETFLKTDYDKCSSVILVPTLFAILNRSEL 300
 QY 303 LDKYDLSNVLVEIASGAPLSKEIGEAVALRRPNLPVROGYGLTETTSAILITPEGDGPK 362
 DB 301 IDKYDLSNVLVEIASGAPLSKEIGEAVALRRPNLPVROGYGLTETTSAILITPEGDGPK 360
 QY 363 ASGVNVLVFAKVIDLDTKKTILGNRRGECVKGPMIMKGYVNDPEATRETIIDEGMLH 422
 DB 361 AVGVNVLVFAKVIDLDTKKTILGNRRGECVKGPMIMKGYVNDPEATRETIIDEGMLH 420
 QY 423 GDIGYDEKHEFFIVDRLSLTKIKYQVPPAELESVILQHPNIFDAGVAPPPPAGE 482
 DB 421 GDIGYDEKHEFFIVDRLSLTKIKYQVPPAELESVILQHPNIFDAGVAPPPPAGE 480
 QY 483 PGAVVVLKKGKSMTEKEMDVYASQVSNARKLRGVRVDEVPKGLTGKIDGAIREILK 542
 DB 481 PAAVVVLEHGKMTKEKTEIVDVAQVTTAKKLRGVVFEVDEVPKGLTGKIDGAIREILK 540
 QY 543 K 543
 DB 541 K 541

RESULT 8
 Q9GPF9 PRELIMINARY; PRT; 548 AA.

AC 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE LUCIFERASE.
 OS Pyrococcus rufa.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidea; Lampyridae; Pyrocoelia.
 OX NCBI_TaxID=71223;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIGHT ORGAN;
 RA Lee K.S., Park H.J., Bae J.S., Lee K.S., Shon H.D., Jin B.R.;

RT "Molecular Cloning and Expression of cDNA encoding luciferase from the
 RT firefly, *Pyrococcus rufa*," to the EMBL/GenBank/DBJ databases.
 RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF328553; AAC45439.1; -
 DR InterPro: IPR000873; AMP-bind.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00284; SERPIN; UNKNOW.
 DR
 SQ SEQUENCE 548 AA; 60787 MW; FE86DEB7DDDEFA9B CRC64;

Query Match 67.3%; Score 1900; DB 5; Length 548;
 Best Local Similarity 66.0%; Pred. No. 5.8e-123;
 Matches 367; Conservative 72; Mismatches 99; Indels 2; Gaps 2;

QY 4 MEND-ENTYGPPEPPYPIEESGAGOLKRYMDRAKL-GAIAFTNALGVDTYAEYLEK 61
 DB 1 MEDDSKHLHGHRSILMEDGTAGEOLHKAKRYAOVGTIAFTDAHEVNTYAEYSEM 60
 QY 62 SCLGEALKNYGLVVDGRIALCSENCEFFIPLVAGLFIGVAVPTNEITYLRELVSIG 121
 DB 61 SCRLAETMKRYGLGQHHIYVCSNSLOFPMVCGALFIGVAVPTNDIYERELYSILF 120
 QY 122 ISKPTIVSSKKGLDKYTVOKTVAITVILDSKVDYRGYQSDNFIKNTPOGFKGSS 181
 DB 121 ISQPTIVSVSKKGLQKILNVQKLPILQKILIMDSKTDYQGFQSMYFVTSHPGCFNEY 180
 QY 182 FKTYEVNRKEQVALIMNSSGSTGLPKGVOLTHENITVTRSHADPIYGNVSPGTALTY 241
 DB 181 DYIPDSFDEETALIMNSSGSTGLPKGVOLTHENITVTRSHADPIYGNVSPGTALTY 240
 QY 242 VPFHNGFMFTTGLYLCGFRIVMLTKFDEETFLKTDYDKCSSVILVPTLFAILNRSEL 301
 DB 241 VPFHNGFMFTTGLYLCGFRIVMLTKFDEETFLKTDYDKCSSVILVPTLFAILNRSEL 300
 QY 302 LDKYDLSNVLVEIASGAPLSKEIGEAVALRRPNLPVROGYGLTETTSAILITPEGDGPK 361
 DB 301 LDKYDLSNVLVEIASGAPLSKEIGEAVALRRPNLPVROGYGLTETTSAILITPEGDGPK 360
 QY 363 ASGVNVLVFAKVIDLDTKKTILGNRRGECVKGPMIMKGYVNDPEATRETIIDEGMLH 421
 DB 361 ASGVNVLVFAKVIDLDTKKTILGNRRGECVKGPMIMKGYVNDPEATRETIIDEGMLH 420
 QY 422 GDIGYDEKHEFFIVDRLSLTKIKYQVPPAELESVILQHPNIFDAGVAPPPPAGE 481
 DB 421 GDIGYDEKHEFFIVDRLSLTKIKYQVPPAELESVILQHPNIFDAGVAPPPPAGE 480
 QY 482 LPAVVVLEHGKMTKEKTEIVDVAQVTTAKKLRGVVFEVDEVPKGLTGKIDGAIREIL 541
 DB 481 LPAVVVLEHGKMTKEKTEIVDVAQVTTAKKLRGVVFEVDEVPKGLTGKIDGAIREIL 540

RESULT 9
 Q26076 PRELIMINARY; PRT; 548 AA.

AC 026076;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE LUCIFERASE.
 OS Pyrococcus miyako.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidea; Lampyridae; Pyrocoelia.
 OX NCBI_TaxID=39363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=WHOLE BODY;
 RX MEDLINE=96061635; PubMed=7480137;
 RA Ohmura Y., Ohba N., Toh H., Tsuji F.I.;
 RT "Cloning, expression and sequence analysis of cDNA for the luciferases
 from the Japanese fireflies, *Pyrococcus miyako* and *Hotaria parvula*,"
 RL Photochem. Photobiol. 62:309-313(1995).

DR EMBL: L39928.1; -
DR HSSP: P08659; ILC1.
DR InterPro: IPR000873; AMP-bind.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00501; AMP-binding_1.
DR PROSITE: PS00455; AMP_BINDING_1.
DR PROSITE: PS00284; SERPIN: UNKNOWN_1.
SO SEQUENCE 548 AA: 60956 MW: CD57F63EE9829401 CRC64

DR EMBL: D25415; BAA05005.1; -
 DR HSSP: P08659; IBA3.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-BINDING; 1.
 SQ SEQUENCE 552 AA; 61000 MW; 85C14ED52BD5366A CRC64;

Query Match 64.6%; Score 1825; DB 5; Length 552;
 Best Local Similarity 62.5%; Pred. No. 8,9e-118;
 Matches 338; Conservative 92; Mismatches 109; Indels 2; Gaps 2;

DB 4 MENDENIVGPEPPFYPIEGSAGAOIRKYMDRYAKL-GAIAFNALIGVDTYAELEKS 62
 1 MSIEENLILGPPYPLIEGTAGEOLHRAITRYAAVGLATYDTLELVYKEFLDY 60
 63 CCIGALKNGLVVDGRICALSCENCEFFIPVLAGLFIGVAVPANEIYTLRELVSLSGI 122
 61 CRALAEAKNGGLGLOHTISVSCENCQVFPVCAALYIGVATPNDIYNERELVSLSI 120
 123 SKPTIVSSKKGLDKYITVQKTYTAIKTYIILDSKYDYGQSDMNFIRKNTPOGKSS 182
 121 SQPTIVVTSRHSLOKILIGVOSRLPVIKIIMLDTKKDYLGQSDMNFIRKNTPOGKSS 180
 183 FKTYEVNRKQVALIMNSGSGTGLPKGVOLTHENIVTRFSHARDPIYGNVSGTALIV 242
 181 FKPLSEF-LDRVACIMNSGSGTGLPKGVPISHRNTTYRSHCHDPYFGNOIIPDITLCA 239
 243 VPFHFGMFTTILGTYLTCGFRIVMLTKFDETEFLKTLQDYKCSSVILPFLALNSEL 302
 240 VPFHAGFTTNGIYIICGRHVYLMVRFNEHLFLQTLQDYKCSSVILPFLALNSEL 299
 303 LDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVQGYGLTETSALITTEGDDKPG 362
 300 VDKYDLHLHEIASGAPLSKEIGEAVARFNLPGVQGYGLTETSALITTEGDDKPG 359
 363 ASGVVPLFKAKYIDLDTKTKLIGPNRGEVCGVPMILMKGYVNDPEATREILIEEGMLHT 422
 360 AVGVVPLFKAKYIDLDTKTKLIGPNRGEVCGVPMILMKGYVNDPEATREILIEEGMLHT 419
 423 GDIGYDEEKEHFTVDRKSLIKYKGYVPPALESVLLQHPNIEDGAVGVPDPIAGEL 482
 420 GDIGYDEEKEHFTVDRKSLIKYKGYVPPALESVLLQHPNIEDGAVGVPDPIAGEL 479
 483 PGAVVVLKKGSMTEKEMVYASOVSNARKLRGCVFVDEYVPGKLGIDKAKIRIELK 542
 480 PGAVVVLKKGSMTEKEMVYASOVSNARKLRGCVFVDEYVPGKLGIDKAKIRIELK 539
 QY 543 K 543
 DB 540 K 540

RESULT 12
 ID 027757 PRELIMINARY; PRT; 545 AA.
 AC 027757; 002653;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).
 OS Photuris pennsylvanica.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Elateriformia; Cantharoidea; Lampyridae; Photuris.
 OC NCBI_TaxID=41716;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LANTERN;
 RX MEDLINE=97307756; PubMed=9165098;
 RA Ye L., Buck L.M., Schaeffer H.J., Leach F.R.;
 RT "Cloning and sequencing of a cDNA for firefly luciferase from Photuris pennsylvanica.";

RL Biochim. Biophys. Acta 1339:39-52(1997).
 CC - FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 562 NM (BY
 CC SIMILARITY).
 CC - CATALYTIC ACTIVITY: PHOTINUS LUCIFERIN + O(2) + ATP - OXIDIZED
 CC PHOTINUS LUCIFERIN + CO(2) + H(2)O + AMP + DIPHOSPHATE + LIGHT.
 CC - COFACTOR: REQUIRES A MAGNESIUM ION (BY SIMILARITY).
 CC - SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
 CC - SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
 DR EMBL: U31240; AAB0897.1; -
 DR HSSP: P08659; ILC1.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-BINDING; UNKNOWN.1.
 KW Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
 KW Peroxisome.
 FT SITE 543 545 MICROBODY TARGETING SIGNAL (POTENTIAL).
 SQ SEQUENCE 545 AA; 60649 MW; F0FE4E828047C26E CRC64;

Query Match 57.6%; Score 1625.5; DB 5; Length 545;
 Best Local Similarity 56.7%; Pred. No. 5.2e-104;
 Matches 308; Conservative 96; Mismatches 138; Indels 1; Gaps 1;

QY 7 DENIVYVPEPPFYPIEGSAGAOIRKYMDRYAKL-GAIAFNALIGVDTYAELEKSCL 65
 DB 3 DKNILIGPEPPFHLADGTAGEOMFALSRADISCIALTNAHTKENVLYEFLKLSGL 62
 QY 66 GEALKNGLVVDGRICALSCENCEFFIPVLAGLFIGVAVPANEIYTLRELVHSIGISKP 125
 DB 63 ABEFKRYGKLDNDTJAVCSENIGLQFLPLASLYGLIINAPVSDYIERELHLSGIVKP 122
 QY 126 TIVSSKKGLDKYITVQKTYTAIKTYIILDSKYDYGQSDMNFIRKNTPOGKSSSEKT 185
 DB 123 RIFGSKNTFOKVLNKSILKTYETIILIDNEDLGQCLNFIQNSNDILDYKKEFP 182
 QY 186 VEVRNRQVALIMNSGSGTGLPKGVOLTHENIVTRFSHARDPIYGNVSGTALIV 245
 DB 183 NSFRDQVALVWFSSGTGVSQVMLTRKNIIVARFNSHCKDPTFGNALNPFTALIV 242
 QY 246 HNGFGFTTILGTYLTCGFRIVMLTKFDETEFLKTLQDYKCSSVILPFLALNSEL 305
 DB 243 HNGFGFTTILGTYLTCGFRIVMLTKFDETEFLKTLQDYKCSSVILPFLALNSEL 302
 QY 306 YDLSNLVEIASGAPLSKEIGEAVARFNLPGVQGYGLTETSALITTEGDDKPGASC 365
 DB 303 YDLSNLVEIASGAPLSKEIGEAVARFNLPGVQGYGLTETSALITTEGDDKPGASC 362
 QY 366 KVPVPLFKAKYIDLDTKTKLIGPNRGEVCGVPMILMKGYVNDPEATREILIEEGMLHT 425
 DB 363 KVPVPLFKAKYIDLDTKTKLIGPNRGEVCGVPMILMKGYVNDPEATREILIEEGMLHT 422
 QY 426 GYDDEEKEHFTVDRKSLIKYKGYVPPALESVLLQHPNIEDGAVGVPDPIAGELPGA 485
 DB 423 GYDDEEKEHFTVDRKSLIKYKGYVPPALESVLLQHPNIEDGAVGVPDPIAGELPGA 482
 QY 486 VVVLKKGSMTEKEMVYASOVSNARKLRGCVFVDEYVPGKLGIDKAKIRIELK 545
 DB 483 GVVVQTKYLINEQVONFVSSQVSTAKMLRGVAKFLDEIPKSGTGKIDRVLMKEMKHK 542
 QY 546 AKM 548
 DB 543 SKL 545

RESULT 13
 ID 090408 PRELIMINARY; PRT; 545 AA.
 AC 090408;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE LUCIFERASE.


```

Db 124 KPIIFASKITIDRVAKVAANKFVKGIALSGSKFFKNYDLKELMEDEKFRQp--- 160
Oy 179 KGSSKRYIVENRKEDVALIMSSGSTGLPKGVOLTHENIYTRSHADPLYGNOVSP--- 235
Db 181 --DFTSPAANKDEBVSLLVCSSGTTGLPKGVOLTOHMLATLD-----SOIQPTVI 229
Oy 236 --GTAILTVPRFHHGFMFTTLGVLNGCFRIWLTGKFDDETFPLTKLODYKCGSSVILVPT 292
Db 230 PHEVTLTLTVIPWHRABSCLLTLTTACYGARLVLLPKFEELKLSAIEKRYVMMAFVWP 289
Oy 293 LFAILNRSLELDKDYLSNLVIEIASGAPLSKEICEIAVARREPLGVNOGGLTETTSAIL 352
Db 290 LMFVLAKHPIDVKYDLSLAWLLCGAPLSRETEDQIKERIGVPIHQVGLSSTLSVL 349
Oy 353 ITPBEDDKPGASGKVVPLFFKAKVLDLDTKTLGNRNGEVCYKGPMLKSGVNDPEATRE 412
Db 350 VQNEBFCKPGSVGLKQVIAKVIDPBTGKLLGNNEGELCFKQDGLMKYVIGDTSQTQ 409
Oy 413 IIDEGCMHTDIDIGYDEDEKHEFFIVDRLSLITKKGQVPAELSEVLLDHNPILFAGVA 472
Db 410 AI-KDGMHTDIDIGYDDDEFPIYDRKLELKIKKGQVPAELTEALLTNDKIKDAVLT 468
Oy 473 GVPPPIAELGAVVVLKKGSKMTEKEMVDVAVASOVNAKRLRGVAFVDEVPKGLTGKI 532
Db 469 GKPEDEAEELPLAVVKQANQVLINENEVIGQVNDNASPARLRGVLTFVDEIKRNSGKI 528
Oy 533 DGKAIRELKRPVAKM 548
Db 529 LRLIREMLKOKSKL 544

```

Search completed: November 29, 2001, 01:59:33
Job time: 378 sec

	Query Match	36.2%	Score 1023;	DB 5:	Length 544;	
	Best Local Similarity	41.5%;	Pred. No. 2, le=62;			
	Matches 231;	Conservative 98;	Mismatches 191;	Indels 36;	Gaps 11	
Oy	9	NIVYGEPEFFPIEBGSA--GAOLRRY-MDRVAKLG-AIAFTNALGVDTVAEYLEKSCC	64			
Db	9	NIVYGG----PYTEROADDSRSLGGLIIDKYKSPFGDRTVLVAAVNGVEYS-ASFMRKSIT	63			
Oy	65	-LGEALIKNYGLVDRIALICSENCEEFFIPVLAGFIGNVGVAPTNITYTLRELTVHSLGIS	123			
Db	64	RLAYILQLKLGVQNNVVYGLSSSENSYNFALAMTAGLAAGTVAPLNTTYLDREYDAHAINIS	123			
Oy	124	KPTLYFSKKGLDKVTYYQKVTYATKAITYILDS-----KVDRYGVSOMDNFKTKNPQGF	178			

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:00:20 ; Search time 26.42 Seconds
(without alignments)
466,760 Million cell updates/sec

Title: US-09-581-241-4

Perfect score: 2823
Sequence: 1 MENMENDENIVGPEPPFYI.....TKIDKAIKRLKRVAKM 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 segs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents-AA:*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTOS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2819	99.9	548	2	US-08-460-934-2
2	2819	99.9	548	2	US-08-782-118-2
3	2817	99.8	548	3	US-09-111-752-14
4	2814	99.7	548	1	US-07-903-047-8
5	2814	99.7	548	4	US-09-380-061B-16
6	2814	99.7	636	2	US-08-460-934-9
7	2814	99.7	636	2	US-08-782-118-9
8	2801	99.2	568	2	US-08-460-934-6
9	2801	99.2	568	2	US-08-782-118-6
10	2703	95.7	552	3	US-09-111-752-10
11	2684	95.1	548	1	US-07-675-211-2
12	2684	95.1	548	1	US-07-903-047-2
13	2684	95.1	548	1	US-08-076-042-2
14	2684	95.1	548	4	US-09-380-061B-14
15	2588	91.7	552	3	US-09-111-752-7
16	2394	84.8	552	3	US-09-111-752-5
17	2335.5	82.7	548	4	US-09-380-061B-18
18	1965.5	69.6	548	4	US-09-380-061B-20
19	1965.5	69.3	550	1	US-08-354-240A-4
20	1947.5	69.0	550	1	US-08-354-240A-6
21	1945.5	68.9	549	1	US-08-354-240A-2
22	1945.5	68.9	550	3	US-08-867-352-23
23	1945.5	68.9	550	4	US-09-380-061B-21
24	1945.5	68.9	815	1	US-08-122-520C-9
25	1942.5	68.8	550	4	US-08-718-425-5
26	1942.5	68.8	550	4	US-08-875-277A-4
27	1939.5	68.7	550	4	US-08-718-425-2

28	1939.5	68.7	550	4	US-08-875-277A-2	Sequence 2, Appli
29	1937.5	68.6	550	4	US-09-380-061B-6	Sequence 6, Appli
30	1905.5	67.5	561	2	US-08-474-169-8	Sequence 8, Appli
31	1831	64.9	552	1	US-08-231-729B-6	Sequence 6, Appli
32	724.5	25.7	540	4	US-08-991-677-8	Sequence 8, Appli
33	585	20.7	578	4	US-08-981-215-1	Sequence 1, Appli
34	258.5	9.2	798	2	US-08-222-617A-8	Sequence 2, Appli
35	256	9.1	582	3	US-08-403-852D-20	Sequence 20, Appli
36	256	9.1	582	3	US-08-510-646B-21	Sequence 21, Appli
37	256	9.1	582	3	US-09-231-818-20	Sequence 20, Appli
38	253	9.0	2618	4	US-09-413-814-28	Sequence 28, Appli
39	251.5	8.9	669	2	US-07-861-800-2	Sequence 2, Appli
40	251	8.9	2628	4	US-09-413-814-11	Sequence 11, Appli
41	243	8.6	3665	2	US-08-222-617A-13	Sequence 13, Appli
42	243	8.6	3712	2	US-08-222-617A-4	Sequence 4, Appli
43	243	8.6	3712	2	US-08-222-617A-25	Sequence 25, Appli
44	231	8.2	768	2	US-08-222-617A-5	Sequence 5, Appli
45	231	8.2	3666	2	US-08-222-617A-12	Sequence 12, Appli

ALIGNMENTS

RESULT 1
US-08-460-934-2
Sequence 2, Application US/08460934
Patent No. 5814465
GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: KIKUCHI, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLO, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460, 934
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLO, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids

TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola lateralis
 US-08-460-934-2

Query Match 99.9%; Score 2819; DB 2; Length 548;
 Best Local Similarity 99.8%; Pred. No. 3.4e-286;
 Matches 547; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNNDENIYGGPEPPPIPIEGSAGAO LRKYMRYAKLGAIAFTNALTGVDYTAAYLE 60
 DB 1 MNNNDENIYGGPEPPPIPIEGSAGAO LRKYMRYAKLGAIAFTNALTGVDYTAAYLE 60
 QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPVLAGLFGVGAPAPNEIYTLRELVHSL 120
 DB 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPVLAGLFGVGAPAPNEIYTLRELVHSL 120
 QY 121 GISKPPIVSSKKGIDVITVOKTVTAIKTIVILDSKVDRYQSDMNFITKKNTPGSGFG 180
 DB 121 GISKPPIVSSKKGIDVITVOKTVTAIKTIVILDSKVDRYQSDMNFITKKNTPGSGFG 180
 QY 181 SSFETVEVNRKEQVALIMNSSSGTGLPKGVOLTHENLVTRFSHARPIYGNVSPGTAIL 240
 DB 181 SSFETVEVNRKEQVALIMNSSSGTGLPKGVOLTHENLVTRFSHARPIYGNVSPGTAIL 240
 QY 241 TVVPFHGFGMFTTGLTGCGRIVMLTKPDEETFKTLTDYKSSVILVPTLFAILNRS 300
 DB 241 TVVPFHGFGMFTTGLTGCGRIVMLTKPDEETFKTLTDYKSSVILVPTLFAILNRS 300
 QY 301 ELIDKYLDSNLVELASGAPLSKEIGAVARRNLPGVRGSGLTETTSAILITPBGDDK 360
 DB 301 ELIDKYLDSNLVELASGAPLSKEIGAVARRNLPGVRGSGLTETTSAILITPBGDDK 360
 QY 361 PGASGVVPLFKAKVIDLDTKTKLGPNRGECVCKPMLMKGVDPNPATREIIDEEGWL 420
 DB 361 PGASGVVPLFKAKVIDLDTKTKLGPNRGECVCKPMLMKGVDPNPATREIIDEEGWL 420
 QY 421 HTGDIGYDEKHFYIDRLKSLIKTKYQVPAELLESVILDHPIFDAGVAGVDPPIAG 480
 DB 421 HTGDIGYDEKHFYIDRLKSLIKTKYQVPAELLESVILDHPIFDAGVAGVDPPIAG 480
 QY 481 ELPGAVVVLKKGKSMTEKEVMYDVASOVNAKRLRGVRFVDEVPKGLTGKIDGAIREI 540
 DB 481 ELPGAVVVLKKGKSMTEKEVMYDVASOVNAKRLRGVRFVDEVPKGLTGKIDGAIREI 540
 QY 541 LKRPVAKM 548
 DB 541 LKRPVAKM 548

RESULT 2

US-08-782-118-2
 Sequence 2, Application US/08782118
 Patent No. 5843746

GENERAL INFORMATION:
 APPLICANT: TATSUMI, HIROKI
 APPLICANT: FUKUDA, SATOSHI
 APPLICANT: KIKUCHI, MAMORU
 APPLICANT: KOYAMA, YASUJI
 TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
 TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
 TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCLELLAND, WATERS & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
 CITY: ARLINGTON

STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/782,118
 FILING DATE: 13-JAN-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/460,934
 FILING DATE: 05-JUN-1995
 APPLICATION NUMBER: JP 193798/1994
 FILING DATE: 27-JUL-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 54625/1995
 FILING DATE: 14-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 98857/1995
 FILING DATE: 24-APR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 7126-001-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 548 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola lateralis
 US-08-782-118-2

Query Match 99.9%; Score 2819; DB 2; Length 548;
 Best Local Similarity 99.8%; Pred. No. 3.4e-286;
 Matches 547; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNNDENIYGGPEPPPIPIEGSAGAO LRKYMRYAKLGAIAFTNALTGVDYTAAYLE 60
 DB 1 MNNNDENIYGGPEPPPIPIEGSAGAO LRKYMRYAKLGAIAFTNALTGVDYTAAYLE 60
 QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPVLAGLFGVGAPAPNEIYTLRELVHSL 120
 DB 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPVLAGLFGVGAPAPNEIYTLRELVHSL 120
 QY 121 GISKPPIVSSKKGIDVITVOKTVTAIKTIVILDSKVDRYQSDMNFITKKNTPGSGFG 180
 DB 121 GISKPPIVSSKKGIDVITVOKTVTAIKTIVILDSKVDRYQSDMNFITKKNTPGSGFG 180
 QY 181 SSFETVEVNRKEQVALIMNSSSGTGLPKGVOLTHENLVTRFSHARPIYGNVSPGTAIL 240
 DB 181 SSFETVEVNRKEQVALIMNSSSGTGLPKGVOLTHENLVTRFSHARPIYGNVSPGTAIL 240
 QY 241 TVVPFHGFGMFTTGLTGCGRIVMLTKPDEETFKTLTDYKSSVILVPTLFAILNRS 300
 DB 241 TVVPFHGFGMFTTGLTGCGRIVMLTKPDEETFKTLTDYKSSVILVPTLFAILNRS 300
 QY 301 ELIDKYLDSNLVELASGAPLSKEIGAVARRNLPGVRGSGLTETTSAILITPBGDDK 360
 DB 301 ELIDKYLDSNLVELASGAPLSKEIGAVARRNLPGVRGSGLTETTSAILITPBGDDK 360
 QY 361 PGASGVVPLFKAKVIDLDTKTKLGPNRGECVCKPMLMKGVDPNPATREIIDEEGWL 420
 DB 361 PGASGVVPLFKAKVIDLDTKTKLGPNRGECVCKPMLMKGVDPNPATREIIDEEGWL 420

QY 421 HTGDIGYDEEKEHFIYDRLSLIKYGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
DB 421 HTGDIGYDEEKEHFIYDRLSLIKYGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
QY 481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNARKLRGCVRFVDEVPKGLTGKIDGKAIRREI 540
DB 481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNARKLRGCVRFVDEVPKGLTGKIDGKAIRREI 540
QY 541 LKKPVAKM 548
DB 541 LKKPVAKM 548

RESULT 3
US-09-111-752-14
; Sequence 14, Application US/09111752
; Patent No. 6074859
; GENERAL INFORMATION:
; APPLICANT: HIRAKAWA, KOZO
; APPLICANT: KAJIYAMA, NAOKI
; APPLICANT: MURAKAMI, SEIJI
; TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
; TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE LUMINESCENT PROTEIN
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,752
; FILING DATE: 08-JUL-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 7126-0009-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-111-752-14

Query Match 99.8%; Score 2817; DB 3; Length 548;
Best Local Similarity 99.8%; Pred. No. 5.5e-286;
Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MENMENDENIVYGPPEFYPIEGSGAGOLRKYMRYAKIGATAFNATLGVDTYVAEYLE 60
DB 1 MENMENDENIVYGPPEFYPIEGSGAGOLRKYMRYAKIGATAFNATLGVDTYVAEYLE 60
QY 61 KSCGCGEALKNGLVWDGRIALSCSECEFFIPVLGLFISGVAPATNEIYTLRELVSIL 120
DB 61 KSCGCGEALKNGLVWDGRIALSCSECEFFIPVLGLFISGVAPATNEIYTLRELVSIL 120
QY 121 GISKPTIVSSKKGLDKVITVOKTVAITKTIYILDSKDYRGYSMDNFIRKNTPOGKRG 180
DB 121 GISKPTIVSSKKGLDKVITVOKTVAITKTIYILDSKDYRGYSMDNFIRKNTPOGKRG 180

DB 121 GISKPTIVSSKKGLDKVITVOKTVAITKTIYILDSKDYRGYSMDNFIRKNTPOGKRG 180
QY 181 SSFKEVNRKROVALIMSSGSTGLPKGVOLTHENLVTRFSHARDPIYGNVSPGTAIL 240
DB 181 SSFKEVNRKROVALIMSSGSTGLPKGVOLTHENLVTRFSHARDPIYGNVSPGTAIL 240
QY 241 TVVFFHGFHGMFTLLGYLTCGFRIVMLTKFDETEFLKTLQDYKCSSVILVPTLAILNRS 300
DB 241 TVVFFHGFHGMFTLLGYLTCGFRIVMLTKFDETEFLKTLQDYKCSSVILVPTLAILNRS 300
QY 301 ELLDKYDLSNLVPIASGAPLSEIGEAVARRPNLPGVNQGIGLETTSAILITTEPGDK 360
DB 301 ELLDKYDLSNLVPIASGAPLSEIGEAVARRPNLPGVNQGIGLETTSAILITTEPGDK 360
QY 361 PGASGVVPLFKAKVIDLTFTKTLGPNRRCGEVCYKGPMLMKGYVDNPEATREIIDEENGL 420
DB 361 PGASGVVPLFKAKVIDLTFTKTLGPNRRCGEVCYKGPMLMKGYVDNPEATREIIDEENGL 420
QY 421 HTGDIGYDEEKEHFIYDRLSLIKYGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
DB 421 HTGDIGYDEEKEHFIYDRLSLIKYGYVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
QY 481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNARKLRGCVRFVDEVPKGLTGKIDGKAIRREI 540
DB 481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNARKLRGCVRFVDEVPKGLTGKIDGKAIRREI 540
QY 541 LKKPVAKM 548
DB 541 LKKPVAKM 548

RESULT 4
US-07-903-047-8
; Sequence 8, Application US/07903047
; Patent No. 529285
; GENERAL INFORMATION:
; APPLICANT: KajiYama, Naoki
; APPLICANT: Nakano, Eichi
; TITLE OF INVENTION: Thermostable luciferase Of Firefly,
; TITLE OF INVENTION: Thermostable luciferase Gene Of Firefly, No. 529285a1 Reco
; TITLE OF INVENTION: DNA, And Process For The Preparation Of Thermostable
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/903,047
; FILING DATE: 19920623
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrick, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7005-048
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 869-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: unknown

MOLECULE TYPE: peptide
us-07-903-047-8

Query Match 99.7%; Score 2814; DB 1; Length 548;
Best Local Similarity 99.6%; Pred. No. 1,1e-285;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MENNENDENIYVGEPEPPYPIEESGAGALRKYMRYAKLGAIAFTNALTGVDYTAAYLE 60
  |||||||
DB 1 MENNENDENIYVGEPEPPYPIEESGAGALRKYMRYAKLGAIAFTNALTGVDYTAAYLE 60
QY 61 KSCCLGALKNYGLVNDGRALALCSENCEEFFIPVLAGLFGVGAAPNEIYTLRELHSL 120
  |||||||
DB 61 KSCCLGALKNYGLVNDGRALALCSENCEEFFIPVLAGLFGVGAAPNEIYTLRELHSL 120
QY 121 GISKPTIVFSSKGLDKVITVOKTVTAIKTIVILDSKVDRGYQSMDFIKKNTPOGFKG 180
  |||||||
DB 121 GISKPTIVFSSKGLDKVITVOKTVTAIKTIVILDSKVDRGYQSMDFIKKNTPOGFKG 180
QY 181 SSFRTVEVNRKEQVALIMNSSGSTGLPKGYQLTHENAVTRSHARDPILGNVSPGTAIL 240
  |||||||
DB 181 SSFRTVEVNRKEQVALIMNSSGSTGLPKGYQLTHENAVTRSHARDPILGNVSPGTAIL 240
QY 241 TVVPFHGEGMFTTIGLTCGFRIVMLTKPDEETFLKTLQDYKSSVILVPTLFAILNRS 300
  |||||||
DB 241 TVVPFHGEGMFTTIGLTCGFRIVMLTKPDEETFLKTLQDYKSSVILVPTLFAILNRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRNLPVROGYGLTETTSAILITPBGDDK 360
  |||||||
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRNLPVROGYGLTETTSAILITPBGDDK 360
QY 361 PGASGVVPLFKAKVIDLDTKKTGLGNRGEVCVKPMLKGYVNDPEATREIIDEEGWL 420
  |||||||
DB 361 PGASGVVPLFKAKVIDLDTKKTGLGNRGEVCVKPMLKGYVNDPEATREIIDEEGWL 420
QY 421 HTGDIGYDEEKHFIVDRILSKIKYGYVPALESVLLQHPNIFDAGVAGVPDPIAG 480
  |||||||
DB 421 HTGDIGYDEEKHFIVDRILSKIKYGYVPALESVLLQHPNIFDAGVAGVPDPIAG 480
QY 481 ELPGAVALVLLKKGKSMTEKEMDYVASOVNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
  |||||||
DB 481 ELPGAVALVLLKKGKSMTEKEMDYVASOVNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
QY 541 LKKPVAKM 548
  |||||||
DB 541 LKKPVAKM 548

```

RESULT 5
us-09-380-061B-16

Sequence 16, Application US/09380061B
Patent No. 6265177

GENERAL INFORMATION:

APPLICANT: SQUIRRELL, DAVID JAMES

WHITE, PETER JOHN

LOWE, CHRISTOPHER ROBIN

MURRAY, JAMES AUGUSTUS HENRY

TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESS: NIXON & VANDERHUYE P. C.

STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/380,061B
FILING DATE: 25-Aug-1999
CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB98/01026

FILING DATE: 7-APR-1998

APPLICATION NUMBER: GB 9707468.8

FILING DATE: 11-APR-1997

ATTORNEY/AGENT INFORMATION:

NAME: SADOFF, B. J.

REGISTRATION NUMBER: 36,663

REFERENCE/DOCKET NUMBER: 124-725

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)816-4000

TELEFAX: (703)816-4100

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 548 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Query Match 99.7%; Score 2814; DB 4; Length 548;
Best Local Similarity 99.6%; Pred. No. 1,1e-285;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MENNENDENIYVGEPEPPYPIEESGAGALRKYMRYAKLGAIAFTNALTGVDYTAAYLE 60
  |||||||
DB 1 MENNENDENIYVGEPEPPYPIEESGAGALRKYMRYAKLGAIAFTNALTGVDYTAAYLE 60
QY 61 KSCCLGALKNYGLVNDGRALALCSENCEEFFIPVLAGLFGVGAAPNEIYTLRELHSL 120
  |||||||
DB 61 KSCCLGALKNYGLVNDGRALALCSENCEEFFIPVLAGLFGVGAAPNEIYTLRELHSL 120
QY 121 GISKPTIVFSSKGLDKVITVOKTVTAIKTIVILDSKVDRGYQSMDFIKKNTPOGFKG 180
  |||||||
DB 121 GISKPTIVFSSKGLDKVITVOKTVTAIKTIVILDSKVDRGYQSMDFIKKNTPOGFKG 180
QY 181 SSFRTVEVNRKEQVALIMNSSGSTGLPKGYQLTHENAVTRSHARDPILGNVSPGTAIL 240
  |||||||
DB 181 SSFRTVEVNRKEQVALIMNSSGSTGLPKGYQLTHENAVTRSHARDPILGNVSPGTAIL 240
QY 241 TVVPFHGEGMFTTIGLTCGFRIVMLTKPDEETFLKTLQDYKSSVILVPTLFAILNRS 300
  |||||||
DB 241 TVVPFHGEGMFTTIGLTCGFRIVMLTKPDEETFLKTLQDYKSSVILVPTLFAILNRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRNLPVROGYGLTETTSAILITPBGDDK 360
  |||||||
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRNLPVROGYGLTETTSAILITPBGDDK 360
QY 361 PGASGVVPLFKAKVIDLDTKKTGLGNRGEVCVKPMLKGYVNDPEATREIIDEEGWL 420
  |||||||
DB 361 PGASGVVPLFKAKVIDLDTKKTGLGNRGEVCVKPMLKGYVNDPEATREIIDEEGWL 420
QY 421 HTGDIGYDEEKHFIVDRILSKIKYGYVPALESVLLQHPNIFDAGVAGVPDPIAG 480
  |||||||
DB 421 HTGDIGYDEEKHFIVDRILSKIKYGYVPALESVLLQHPNIFDAGVAGVPDPIAG 480
QY 481 ELPGAVALVLLKKGKSMTEKEMDYVASOVNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
  |||||||
DB 481 ELPGAVALVLLKKGKSMTEKEMDYVASOVNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
QY 541 LKKPVAKM 548
  |||||||
DB 541 LKKPVAKM 548

```

RESULT 6

us-08-460-934-9
Sequence 9, Application US/08460934

Patent No. 5814465
GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: KOYAMA, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINATED AND A BIOLUMINESCENT
TITLE OF INVENTION: ANALYSIS METHOD
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,934
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ. ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-934-9

Query Match 99.7%; Score 2814; DB 2; Length 636;
Best local similarity 99.8%; Pred. No. 1,5e-285;
Matches 546; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEMEMDENIVVGPPEFYPIEBSGAGOLRKYMDRYAKGALAFNALGVGYTAEYLE 60
DB 1 MEMEMDENIVVGPPEFYPIEBSGAGOLRKYMDRYAKGALAFNALGVGYTAEYLE 60
QY 61 KSCCLGALKNYLVVDGRIALSCENCEFFIPLVLAFLIGVAVAPTEIYLRLEIYHSL 120
DB 61 KSCCLGALKNYLVVDGRIALSCENCEFFIPLVLAFLIGVAVAPTEIYLRLEIYHSL 120
QY 121 GISKPTIVSSKKGDKLVTVQKTVTAITVYILDSKVYRGYQSDNFIKKNTPOGFG 180
DB 121 GISKPTIVSSKKGDKLVTVQKTVTAITVYILDSKVYRGYQSDNFIKKNTPOGFG 180
QY 181 SSKTYEVNKKRQVALIMNSGSGTGLPKGVOLTHENLVTRFSHARDPIYGNVSPETAIL 240
DB 181 SSKTYEVNKKRQVALIMNSGSGTGLPKGVOLTHENLVTRFSHARDPIYGNVSPETAIL 240

QY 241 TVPFFHGFQMTTGLTGLTGFRIYMLTKFDEETFLTKLDYKCSSYIIVPTLFAILNRS 300
DB 241 TVPFFHGFQMTTGLTGLTGFRIYMLTKFDEETFLTKLDYKCSSYIIVPTLFAILNRS 300
QY 301 ELLDKYDLSNVEIASGAPLSKEIGAVARRENLPGVROGYGLTETTSAILITTPGDDK 360
DB 301 ELLDKYDLSNVEIASGAPLSKEIGAVARRENLPGVROGYGLTETTSAILITTPGDDK 360
QY 361 PGASGVVPLFKAVVIDLTKTGLPBRGECVCKGMLKGVNDNEATREIIDEIGWL 420
DB 361 PGASGVVPLFKAVVIDLTKTGLPBRGECVCKGMLKGVNDNEATREIIDEIGWL 420
QY 421 HTGDIGYDEKHFIYDRLSLTKYKGYVPPAELESVLLQHPNIEDAGVAGVPPDIAG 480
DB 421 HTGDIGYDEKHFIYDRLSLTKYKGYVPPAELESVLLQHPNIEDAGVAGVPPDIAG 480
QY 481 ELPGAIVVLLKKGKSMTEKEVMDIYASOVSNARKRGVRRVDEVPKGLTGKIDGKAIREI 540
DB 481 ELPGAIVVLLKKGKSMTEKEVMDIYASOVSNARKRGVRRVDEVPKGLTGKIDGKAIREI 540
QY 541 LKKPVAK 547
DB 541 LKKPVAK 547

RESULT 7
US-08-782-118-9
Sequence 9, Application US/08782118
Patent No. 5843746
GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: KOYAMA, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINATED AND A BIOLUMINESCENT
TITLE OF INVENTION: ANALYSIS METHOD
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/782,118
FILING DATE: 13-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,934
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 636 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-782-118-9

Query Match 99.7%; Score 2814; DB 2; Length 636;
 Best local Similarity 99.8%; Pred. No. 1.5e-285;
 Matches 546; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNENDENIYVGEPEFYPIEGSGAQLRKRYMDRYAKLGAIAFTNALTGVDYTAETLE 60
 DB 1 MNNENDENIYVGEPEFYPIEGSGAQLRKRYMDRYAKLGAIAFTNALTGVDYTAETLE 60
 QY 61 KSCCLGALKNYGLVVGRIALACSENGEPEFIPVLGLFIVGVAAPINEIYTLRELVHSL 120
 DB 61 KSCCLGALKNYGLVVGRIALACSENGEPEFIPVLGLFIVGVAAPINEIYTLRELVHSL 120
 QY 121 GISKPTIVSSKGLDVKVTYVOKTVAIKTIVILDSKYDYGQSMDFIKKNTPOGFKG 180
 DB 121 GISKPTIVSSKGLDVKVTYVOKTVAIKTIVILDSKYDYGQSMDFIKKNTPOGFKG 180
 QY 181 SSFRTVEVNRKEOVALIMNSSGSTGLPKGVOLTHEMLVTRFSHARDPIYGNQVSPGTAIL 240
 DB 181 SSFRTVEVNRKEOVALIMNSSGSTGLPKGVOLTHEMLVTRFSHARDPIYGNQVSPGTAIL 240
 QY 241 TVPFHGFMTTGLTGTGFRIVMLTKFDEETFLKTDYKSSVILVPTLAILNRS 300
 DB 241 TVPFHGFMTTGLTGTGFRIVMLTKFDEETFLKTDYKSSVILVPTLAILNRS 300
 QY 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETSAILIIRPEGDDK 360
 DB 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETSAILIIRPEGDDK 360
 QY 361 PGASGRVPLFKAKVIDLDTKTKTGPNRGEVCYKGMKGVYDNEATREIIDEESGWL 420
 DB 361 PGASGRVPLFKAKVIDLDTKTKTGPNRGEVCYKGMKGVYDNEATREIIDEESGWL 420
 QY 421 HTDDIGVDEEKHFETIDRLKSLIKYGYOVPALESVLLQHPNIFDAGVAGVDDIAG 480
 DB 421 HTDDIGVDEEKHFETIDRLKSLIKYGYOVPALESVLLQHPNIFDAGVAGVDDIAG 480
 QY 481 ELPGAVVILKKGSMTEKEVMDYVAVASQVSNAKRLRGVRVDEVPKGLTKIDGKAIREI 540
 DB 481 ELPGAVVILKKGSMTEKEVMDYVAVASQVSNAKRLRGVRVDEVPKGLTKIDGKAIREI 540
 QY 541 LKRPVAK 547
 DB 541 LKRPVAK 547

RESULT 8
 US-08-460-934-6
 Sequence 6, Application US/08460934
 Patent No. 5814465
 GENERAL INFORMATION:
 APPLICANT: TATSUMI, HIROKI
 APPLICANT: FUKUDA, SATOSHI
 APPLICANT: KIKUCHI, MAMORU
 APPLICANT: KOTAMA, YASUJI
 TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
 TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
 TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR

CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/460,934
 FILING DATE: 05-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 193798/1994
 FILING DATE: 27-JUL-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 54625/1995
 FILING DATE: 14-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 98857/1995
 FILING DATE: 24-APR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 7126-001-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 568 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-460-934-6

Query Match 99.2%; Score 2801; DB 2; Length 568;
 Best local Similarity 99.5%; Pred. No. 2.8e-284;
 Matches 543; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 MNNENDENIYVGEPEFYPIEGSGAQLRKRYMDRYAKLGAIAFTNALTGVDYTAETLEKS 62
 DB 23 SLENDENIYVGEPEFYPIEGSGAQLRKRYMDRYAKLGAIAFTNALTGVDYTAETLEKS 82
 QY 63 CCLGEALKNYGLVVDGRIALACSENGEPEFIPVLGLFIVGVAAPINEIYTLRELVHSLGI 122
 DB 83 CCLGEALKNYGLVVDGRIALACSENGEPEFIPVLGLFIVGVAAPINEIYTLRELVHSLGI 142
 QY 123 SKPTIVSSKGLDVKVTYVOKTVAIKTIVILDSKYDYGQSMDFIKKNTPOGFKGSS 182
 DB 143 SKPTIVSSKGLDVKVTYVOKTVAIKTIVILDSKYDYGQSMDFIKKNTPOGFKGSS 202
 QY 183 EKTVEVNRKEOVALIMNSSGSTGLPKGVOLTHEMLVTRFSHARDPIYGNQVSPGTAILTV 242
 DB 203 EKTVEVNRKEOVALIMNSSGSTGLPKGVOLTHEMLVTRFSHARDPIYGNQVSPGTAILTV 262
 QY 243 VPFHGFMTTGLTGTGFRIVMLTKFDEETFLKTDYKSSVILVPTLAILNRSSEL 302
 DB 263 VPFHGFMTTGLTGTGFRIVMLTKFDEETFLKTDYKSSVILVPTLAILNRSSEL 322
 QY 303 LDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETSAILIIRPEGDDKPG 362
 DB 323 LDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETSAILIIRPEGDDKPG 382
 QY 363 ASGRVPLFKAKVIDLDTKTKTGPNRGEVCYKGMKGVYDNEATREIIDEESGWLHT 422
 DB 383 ASGRVPLFKAKVIDLDTKTKTGPNRGEVCYKGMKGVYDNEATREIIDEESGWLHT 442
 QY 423 GDIGYDEEKHFETIDRLKSLIKYGYOVPALESVLLQHPNIFDAGVAGVDDIAGEL 482
 DB 443 GDIGYDEEKHFETIDRLKSLIKYGYOVPALESVLLQHPNIFDAGVAGVDDIAGEL 502

QY 483 PGAVVVLKKGKSMTEKEVMDYVAVASQVSNKRLRGVRFVDEVPKGLTGKIDGKAIRELTK 542
DB 503 PGAVVVLKKGKSMTEKEVMDYVAVASQVSNKRLRGVRFVDEVPKGLTGKIDGKAIRELTK 562
QY 543 KPVAKM 548
DB 563 KPVAKM 568

RESULT 9

US-08-782-118-6
Sequence 6, Application US/08782118
Patent No. 5843746

GENERAL INFORMATION:

APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: KOYAMA, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
TITLE OF INVENTION: ANALYSIS METHOD
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/782,118
FILING DATE: 13-JAN-1997
CLASSIFICATION: 435

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/460,934
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995

ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 568 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-782-118-6

Query Match 99.2%; Score 2801; DB 2; Length 568;

Best Local Similarity 99.5%; Pred. No. 2, 8e-284;
Matches 543; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 NMENDENIVGPEPFYPIEESAGAOQLRKYMRYAKIGAIATFNALTGVDTYAEYLEKS 62

DB 23 SLENDENIVGPEPFYPIEESAGAOQLRKYMRYAKIGAIATFNALTGVDTYAEYLEKS 82
QY 63 CCLDEALKNGLVYDGRALALSCENCEEFFIPVLAGLFIGVAVAPNTNITLRLVSLGI 122
DB 83 CCLDEALKNGLVYDGRALALSCENCEEFFIPVLAGLFIGVAVAPNTNITLRLVSLGI 142
QY 123 SKPTIVSSKKGLDKVITVQKVTVAITVILDSKVYRGYQSMDFIKKNTPOGFSS 182
DB 143 SKPTIVSSKKGLDKVITVQKVTVAITVILDSKVYRGYQSMDFIKKNTPOGFSS 202
QY 183 FKTEVNRKEQVALINSSGSLPKGVOLTHENLVTRFSHARDPIYGNVSPGATLTV 242
DB 203 FKTEVNRKEQVALINSSGSLPKGVOLTHENLVTRFSHARDPIYGNVSPGATLTV 262
QY 243 VPFHFGMFTTGLYLCGRFVWLTFDEETPLKTDYKCSSVILVPLFALNSEL 302
DB 263 VPFHFGMFTTGLYLCGRFVWLTFDEETPLKTDYKCSSVILVPLFALNSEL 322
QY 303 LDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVRGGLTFETTSAILITPEGDDPKG 362
DB 323 LDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVRGGLTFETTSAILITPEGDDPKG 382
QY 363 ASGKVPFLKAKVIDLTKTGLGNRRGVCVKPMLDKGYDNPETREIIDEENLMT 422
DB 383 ASGKVPFLKAKVIDLTKTGLGNRRGVCVKPMLDKGYDNPETREIIDEENLMT 442
QY 423 GDIGYDEKHFIVDRKSLIKKYGQVPALESVLLQHPNIFDGVAGVPPINGEL 482
DB 443 GDIGYDEKHFIVDRKSLIKKYGQVPALESVLLQHPNIFDGVAGVPPINGEL 502
QY 483 PGAVVVLKKGKSMTEKEVMDYVAVASQVSNKRLRGVRFVDEVPKGLTGKIDGKAIRELTK 542
DB 503 PGAVVVLKKGKSMTEKEVMDYVAVASQVSNKRLRGVRFVDEVPKGLTGKIDGKAIRELTK 562
QY 543 KPVAKM 548
DB 563 KPVAKM 568

RESULT 10
US-09-111-752-10
Sequence 10, Application US/09111752
Patent No. 6074859

GENERAL INFORMATION:

APPLICANT: HIROKAWA, KOZO
APPLICANT: KAJIYAMA, NAOKI
APPLICANT: MURAKAMI, SEIJI
TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN AND
PROCESS FOR PRODUCING MUTANT-TYPE BIOLUMINESCENT PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,752
FILING DATE: 08-JUL-1998
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-0009-0
TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 552 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola lateralis, Photinus pyralis
 US-09-111-752-10

Query Match 95.7%; Score 2703; DB 3; Length 552;
 Best Local Similarity 96.3%; Pred. No. 4.8e-274;
 Matches 523; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 1 MENNENDENIYGEPPYPIEESGAGQAKRYMDRYAKLGAIAFTNALTGVDYTAAYLE 60
 DB 1 MENNENDENIYGEPPYPIEESGAGQAKRYMDRYAKLGAIAFTNALTGVDYTAAYLE 60
 QY 61 KSCCLGALNKGVLVNGRIALCSENCEEFFIPVLAGLFGVGAAPNEITYTLREIVHSL 120
 DB 61 KSCCLGALNKGVLVNGRIALCSENCEEFFIPVLAGLFGVGAAPNEITYTLREIVHSL 120
 QY 121 GISKPITVSSKGLDKVITVQKTVTAIKTIVILDSKVDRGYOSMDFIKKNTPOGFKG 180
 DB 121 GISKPITVSSKGLDKVITVQKTVTAIKTIVILDSKVDRGYOSMDFIKKNTPOGFKG 180
 QY 181 SSFRTYVNRKEQVALIMNSSGSTGLPKGVQLTHENLVTRFSHARDPIYGNVSPGTAIL 240
 DB 181 SSFRTYVNRKEQVALIMNSSGSTGLPKGVQLTHENLVTRFSHARDPIYGNVSPGTAIL 240
 QY 241 TVPFHNGFGMFTTIGLTCGFRIVMLTKFDEETFLKTLQDYKSSVILVPTLFAILNRS 300
 DB 241 TVPFHNGFGMFTTIGLTCGFRIVMLTKFDEETFLKTLQDYKSSVILVPTLFAILNRS 300
 QY 301 ELLDKYDLSMLVEIASGAPLSKEIGAVARFNLPGVRGQYGLTETTSAILITPEGDDK 360
 DB 301 ELLDKYDLSMLVEIASGAPLSKEIGAVARFNLPGVRGQYGLTETTSAILITPEGDDK 360
 QY 361 PGASGVKVPLEFAKVIDLDLTKTLGPRRGEVCYKGMKKGVDNDEARREIIDEWGWL 420
 DB 361 PGASGVKVPLEFAKVIDLDLTKTLGPRRGEVCYKGMKKGVDNDEARREIIDEWGWL 420
 QY 421 HTGDIGYDEEKHFPIVDRLSLTKYKGYOVPAALESVLLQHPNIFDAGVAGVPPDIAG 480
 DB 421 HTGDIGYDEEKHFPIVDRLSLTKYKGYOVPAALESVLLQHPNIFDAGVAGVPPDIAG 480
 QY 481 ELPGAAYVILKKGSMTKEKMDYVASOVNAKRLRGVRFVDEVPKGLTGKIDGAIAREI 540
 DB 481 ELPGAAYVILKKGSMTKEKMDYVASOVNAKRLRGVRFVDEVPKGLTGKIDGAIAREI 540
 QY 541 LKK 543
 DB 541 LKK 543

RESULT 11
 US-07-675-211-2
 ; Sequence 2, Application US/07675211
 ; Patent No. 5219737
 ; GENERAL INFORMATION:
 ; APPLICANT: KAITIYAMA, MOKI
 ; APPLICANT: NAANO, EICHI
 ; TITLE OF INVENTION: MUTANT LUCIFERASE OF A FIREFLY, MUTANT
 ; TITLE OF INVENTION: LUCIFERASE GENES, NOVEL RECOMBINANT DNAs CONTAINING THE
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PENNIE & EDMONDS
 ; STREET: 1155 AVENUE OF THE AMERICAS

CITY: NEW YORK
 STATE: N.Y.
 COUNTRY: U.S.A
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/675,211
 FILING DATE: 19910326
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MISROCK, S. LESLIE
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7005-026-999
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 548 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola cruciata
 US-07-675-211-2

Query Match 95.1%; Score 2684; DB 1; Length 548;
 Best Local Similarity 93.4%; Pred. No. 4.6e-272;
 Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

QY 1 MENNENDENIYGEPPYPIEESGAGQAKRYMDRYAKLGAIAFTNALTGVDYTAAYLE 60
 DB 1 MENNENDENIYGEPPYPIEESGAGQAKRYMDRYAKLGAIAFTNALTGVDYTAAYLE 60
 QY 61 KSCCLGALNKGVLVNGRIALCSENCEEFFIPVLAGLFGVGAAPNEITYTLREIVHSL 120
 DB 61 KSCCLGALNKGVLVNGRIALCSENCEEFFIPVLAGLFGVGAAPNEITYTLREIVHSL 120
 QY 121 GISKPITVSSKGLDKVITVQKTVTAIKTIVILDSKVDRGYOSMDFIKKNTPOGFKG 180
 DB 121 GISKPITVSSKGLDKVITVQKTVTAIKTIVILDSKVDRGYOSMDFIKKNTPOGFKG 180
 QY 181 SSFRTYVNRKEQVALIMNSSGSTGLPKGVQLTHENLVTRFSHARDPIYGNVSPGTAIL 240
 DB 181 SSFRTYVNRKEQVALIMNSSGSTGLPKGVQLTHENLVTRFSHARDPIYGNVSPGTAIL 240
 QY 241 TVPFHNGFGMFTTIGLTCGFRIVMLTKFDEETFLKTLQDYKSSVILVPTLFAILNRS 300
 DB 241 TVPFHNGFGMFTTIGLTCGFRIVMLTKFDEETFLKTLQDYKSSVILVPTLFAILNRS 300
 QY 301 ELLDKYDLSMLVEIASGAPLSKEIGAVARFNLPGVRGQYGLTETTSAILITPEGDDK 360
 DB 301 ELLDKYDLSMLVEIASGAPLSKEIGAVARFNLPGVRGQYGLTETTSAILITPEGDDK 360
 QY 361 PGASGVKVPLEFAKVIDLDLTKTLGPRRGEVCYKGMKKGVDNDEARREIIDEWGWL 420
 DB 361 PGASGVKVPLEFAKVIDLDLTKTLGPRRGEVCYKGMKKGVDNDEARREIIDEWGWL 420
 QY 421 HTGDIGYDEEKHFPIVDRLSLTKYKGYOVPAALESVLLQHPNIFDAGVAGVPPDIAG 480
 DB 421 HTGDIGYDEEKHFPIVDRLSLTKYKGYOVPAALESVLLQHPNIFDAGVAGVPPDIAG 480
 QY 481 ELPGAAYVILKKGSMTKEKMDYVASOVNAKRLRGVRFVDEVPKGLTGKIDGAIAREI 540
 DB 481 ELPGAAYVILKKGSMTKEKMDYVASOVNAKRLRGVRFVDEVPKGLTGKIDGAIAREI 540
 QY 541 LKKPVAKM 548

Db 541 LKPYAKM 548

RESULT 12

US-07-903-047-2
; Sequence 2, Application US/07903047
; Patent No. 5229285
; GENERAL INFORMATION:
; APPLICANT: Kajiyama, Naoki
; APPLICANT: Nakano, Eiichi
; TITLE OF INVENTION: Thermostable luciferase of firefly,
; TITLE OF INVENTION: Thermostable luciferase gene of firefly, No. 5229285el Recombi
; TITLE OF INVENTION: DNA, And Process For The Preparation Of Thermostable
; TITLE OF INVENTION: Luciferase of firefly
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/903,047
; FILING DATE: 19920623
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7005-048
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-07-903-047-2

Query Match 95.1%; Score 2684; DB 1; Length 548;
Best Local Similarity 93.4%; Pred. No. 4.6e-272;
Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

QY 1 MEMNENDENIVGPPEFYPIEESGAGQLRKKYMDRYAKLGAIAFTNALTGVDTVAEYLE 60
Db 1 MEMNENDENIVGPPEFYPIEESGAGQLRKKYMDRYAKLGAIAFTNALTGVDTVAEYLE 60
QY 61 KSCCLGALKNYGLVVDGRALSCENCEFFIVLAGLPFGVAVATNNEITRELYVHSL 120
Db 61 KSCCLGALKNYGLVVDGRALSCENCEFFIVLAGLPFGVAVATNNEITRELYVHSL 120
QY 121 GISKPTIVSSKKGDKVITVQKTAIKTIVILDSKVDYRGQSDMNFIKNTPOGFGK 180
Db 121 GISKPTIVSSKKGDKVITVQKTAIKTIVILDSKVDYRGQSDMNFIKNTPOGFGK 180
QY 181 SSFRTVEVNRKEQVALIMSSGSTGLPKGVOLTHENLVTRFSHARDPIYGNVSPCTAVI 240
Db 181 SSFRTVEVNRKEQVALIMSSGSTGLPKGVOLTHENLVTRFSHARDPIYGNVSPCTAVI 240
QY 241 TVVPEHHGFGMTTGLGICGFRVYMLTKPDEETPLTKLQDYKCTSVILVPTLAILNKS 300
Db 241 TVVPEHHGFGMTTGLGICGFRVYMLTKPDEETPLTKLQDYKCTSVILVPTLAILNKS 300

QY 301 ELLDKYDLSNIVETASGAPLSKEIGFAVARREMLPGVROGYGLTETTSAILITTPGDDK 360
Db 301 ELLDKYDLSNIVETASGAPLSKEIGFAVARREMLPGVROGYGLTETTSAILITTPGDDK 360
QY 361 PGASGVVPLFEKAKVIDLDRKTLGPNRGEVGVCKGPMKKGVDNDEATREIIDEQWL 420
Db 361 PGASGVVPLFEKAKVIDLDRKTLGPNRGEVGVCKGPMKKGVDNDEATREIIDEQWL 420
QY 421 HTGIGIYDEEKHFIVDRKLSLIKRYGYVPAELESYLLQHPNIFDAGVAGPPPIAG 480
Db 421 HTGIGIYDEEKHFIVDRKLSLIKRYGYVPAELESYLLQHPNIFDAGVAGPPPIAG 480
QY 481 ELPGAVVLLKKGKSMTEKVDYASOVNAKRLRGVRFVDEVPKGLTGKIDGAKIREI 540
Db 481 ELPGAVVLLKKGKSMTEKVDYASOVNAKRLRGVRFVDEVPKGLTGKIDGAKIREI 540
QY 541 LKPYAKM 548
Db 541 LKPYAKM 548

RESULT 13

US-08-076-042-2
; Sequence 2, Application US/08076042
; Patent No. 5330906
; GENERAL INFORMATION:
; APPLICANT: KAJIYAMA, NAOKI
; APPLICANT: NAKANO, EIICHI
; TITLE OF INVENTION: MUTANT LUCIFERASE OF A FIREFLY, MUTANT
; TITLE OF INVENTION: LUCIFERASE GENES, NOVEL RECOMBINANT DNAs CONTAINING THE
; TITLE OF INVENTION: GENES AND A METHOD OF PRODUCING MUTANT LUCIFERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: N.Y.
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/076,042
; FILING DATE: 15-JUN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/675,211
; FILING DATE: 26-MAR-1991
; NAME: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISTROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7005-026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Luciola cruciata
; US-08-076-042-2

Query Match 95.1%; Score 2684; DB 1; Length 548;

Best Local Similarity 93.4%; Pred. No. 4.6e-272;
Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

```

QY 1 MENNENDENIYVGPPEPIEGSAGAO LRKYMRYAKGIAFTNALTGVDYTAETLE 60
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 1 MENNENDENIYVGPPEPIEGSAGAO LRKYMRYAKGIAFTNALTGVDYTAETLE 60
QY 61 KSCCLGKALQNYGLVVDGRALCSENCEEFPIVYAGLFGVGAAPNEIYTLRELVS 120
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 61 KSCCLGKALQNYGLVVDGRALCSENCEEFPIVYAGLFGVGAAPNEIYTLRELVS 120
QY 121 GISKPTIVFSSKGLDKVITVQKTVTAIKTVIILDSKVDYGYOSMDNFIRKNTPOG 180
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 121 GISKPTIVFSSKGLDKVITVQKTVTAIKTVIILDSKVDYGYOSMDNFIRKNTPOG 180
QY 181 SSFETVEYNKREQVALIMNSSGSTGLPKGYOLTHEINTVTRSHARDPIYGNQVSPGTA 240
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 181 SSFETVEYNKREQVALIMNSSGSTGLPKGYOLTHEINTVTRSHARDPIYGNQVSPGTA 240
QY 241 TVPFHHGFGFTLLGYLTGCFRIYMLTKDEEFTLKTLDYKCSSVILVPTLFAILNRS 300
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 241 TVPFHHGFGFTLLGYLTGCFRIYMLTKDEEFTLKTLDYKCSSVILVPTLFAILNRS 300
QY 301 ELLDKYDLSNVLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETTSAILITPEGDDK 360
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 301 ELLDKYDLSNVLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETTSAILITPEGDDK 360
QY 361 PGASGVVPLFKAKVIDLDTKRTGLPNRGEVCYKGMKGYVNDPEATREIIDEBSGL 420
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 361 PGASGVVPLFKAKVIDLDTKRTGLPNRGEVCYKGMKGYVNDPEATREIIDEBSGL 420
QY 421 HTGDIGYDEKHFETIDRLKSLIKYGYOVPPELESVLLQHNIPIDAGVAGVDPDIAG 480
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 421 HTGDIGYDEKHFETIDRLKSLIKYGYOVPPELESVLLQHNIPIDAGVAGVDPDIAG 480
QY 481 ELPGAIVVLLKKGSMTEKEVMDVYASOVSNARKLRGVRVDEVPKGLTGIDGAI 540
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 481 ELPGAIVVLLKKGSMTEKEVMDVYASOVSNARKLRGVRVDEVPKGLTGIDGAI 540
QY 541 LKRPVAKM 548
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 541 LKRPVAKM 548

```

RESULT 14
US-09-380-061B-14
Sequence 14, Application US/09380061B
Patent No. 6265177
GENERAL INFORMATION:
APPLICANT: SOURRELL, DAVID JAMES
WHITE, PETER JOHN
MURRAY, JAMES AUGUSTUS HENRY
TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHIVE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/380,061B
FILING DATE: 25-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB98/01026

```

; FILING DATE: 7-Apr-1998
; APPLICATION NUMBER: GB 9707468.8
; FILING DATE: 11-Apr-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B. J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 124-725
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4000
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-380-061B-14

```

Query Match 95.1%; Score 2684; DB 4; Length 548;
Best Local Similarity 93.4%; Pred. No. 4.6e-272;
Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

```

QY 1 MENNENDENIYVGPPEPIEGSAGAO LRKYMRYAKGIAFTNALTGVDYTAETLE 60
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 1 MENNENDENIYVGPPEPIEGSAGAO LRKYMRYAKGIAFTNALTGVDYTAETLE 60
QY 61 KSCCLGKALQNYGLVVDGRALCSENCEEFPIVYAGLFGVGAAPNEIYTLRELVS 120
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 61 KSCCLGKALQNYGLVVDGRALCSENCEEFPIVYAGLFGVGAAPNEIYTLRELVS 120
QY 121 GISKPTIVFSSKGLDKVITVQKTVTAIKTVIILDSKVDYGYOSMDNFIRKNTPOG 180
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 121 GISKPTIVFSSKGLDKVITVQKTVTAIKTVIILDSKVDYGYOSMDNFIRKNTPOG 180
QY 181 SSFETVEYNKREQVALIMNSSGSTGLPKGYOLTHEINTVTRSHARDPIYGNQVSPGTA 240
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 181 SSFETVEYNKREQVALIMNSSGSTGLPKGYOLTHEINTVTRSHARDPIYGNQVSPGTA 240
QY 241 TVPFHHGFGFTLLGYLTGCFRIYMLTKDEEFTLKTLDYKCSSVILVPTLFAILNRS 300
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 241 TVPFHHGFGFTLLGYLTGCFRIYMLTKDEEFTLKTLDYKCSSVILVPTLFAILNRS 300
QY 301 ELLDKYDLSNVLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETTSAILITPEGDDK 360
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 301 ELLDKYDLSNVLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETTSAILITPEGDDK 360
QY 361 PGASGVVPLFKAKVIDLDTKRTGLPNRGEVCYKGMKGYVNDPEATREIIDEBSGL 420
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 361 PGASGVVPLFKAKVIDLDTKRTGLPNRGEVCYKGMKGYVNDPEATREIIDEBSGL 420
QY 421 HTGDIGYDEKHFETIDRLKSLIKYGYOVPPELESVLLQHNIPIDAGVAGVDPDIAG 480
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 421 HTGDIGYDEKHFETIDRLKSLIKYGYOVPPELESVLLQHNIPIDAGVAGVDPDIAG 480
QY 481 ELPGAIVVLLKKGSMTEKEVMDVYASOVSNARKLRGVRVDEVPKGLTGIDGAI 540
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 481 ELPGAIVVLLKKGSMTEKEVMDVYASOVSNARKLRGVRVDEVPKGLTGIDGAI 540
QY 541 LKRPVAKM 548
    ||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 541 LKRPVAKM 548

```

RESULT 15
US-09-111-752-7
Sequence 7, Application US/09111752
Patent No. 6074859
GENERAL INFORMATION:
APPLICANT: HIROKAWA, KOZO
APPLICANT: KAJIYAMA, NAOKI
APPLICANT: MURAKAMI, SEIJI

;; TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
;; TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE BIOLUMINESCENT PROTEIN
;; NUMBER OF SEQUENCES: 14
;; CORRESPONDENCE ADDRESS:
;; ADDRESS: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;; ADDRESS: P. C. JEFFERSON DAVIS HIGHWAY, SUITE 400
;; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;; CITY: ARLINGTON
;; STATE: VA
;; COUNTRY: USA
;; ZIP: 22202
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/111,752
;; FILING DATE: 08-JUL-1998
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: OBLON, NORMAN F.
;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 7126-0009-0
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-413-3000
;; TELEFAX: 703-413-2220
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 552 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; ORIGINAL SOURCE:
;; ORGANISM: Luciola cruciata and Photinus pyralis
;;
;; US-09-111-752-7

Query Match 91.7%; Score 2588; DB 3; Length 552;
Best Local Similarity 90.8%; Pred. No. 5.3e-262;
Matches 493; Conservative 32; Mismatches 18; Indels 0; Gaps 0;

QY 1 MNNNDNENIYVGPPEPIPIEGSSAGAOIKRYMDYAKGAIAPTNALTGVDYTYAEYLE 60
DB 1 MNNNDNENIYVGPPEPIPIEGSSAGAOIKRYMDYAKGAIAPTNALTGVDYTYAEYLE 60
QY 61 KSCCLGKALKNYGLVVDGRIALCSENCPEPIVLAGLFTGVGVAPTNETYTLRELHVS 120
DB 61 KSCCLGKALKNYGLVVDGRIALCSENCPEPIVLAGLFTGVGVAPTNETYTLRELHVS 120
QY 121 GISKPTIVFSSKGLDKVITVQKTIVTAIKTIVILDSKVYKGYQSMDFIKKNTPOGFKG 180
DB 121 GISKPTIVFSSKGLDKVITVQKTIVTAIKTIVILDSKVYKGYQSMDFIKKNTPOGFKG 180
QY 181 SSFETVENRKREQVALLMNSSGSTGLPKGYQLTHENLVTRSHARDPYGNQVSPGTAIL 240
DB 181 SSFETVENRKREQVALLMNSSGSTGLPKGYQLTHENLVTRSHARDPYGNQVSPGTAIL 240
QY 241 TVVPFHGFGMTTLGLYTCGFRIVMLTKFDEETFLKTLDYKCSSVILVPTLFAILLNS 300
DB 241 TVVPFHGFGMTTLGLYTCGFRIVMLTKFDEETFLKTLDYKCSSVILVPTLFAILLNS 300
QY 301 ELDKYLDSLNLVETASGAPLSKEIGEAVARFNLPGVRQYGLTETTSAILITPEGDDK 360
DB 301 ELDKYLDSLNLVETASGAPLSKEIGEAVARFNLPGVRQYGLTETTSAILITPEGDDK 360
QY 361 PGAGKVVPLFKAKVIDLDTKKTLGPNRSGEYCYKPMILMKGYVNDPEATREITIDEGWL 420
DB 361 PGAGKVVPLFKAKVIDLDTKKTLGPNRSGEYCYKPMILMKGYVNDPEATREITIDEGWL 420
QY 421 HTGDIGYYDEEKHFFIVDRILKSLIKYGYOVPPALESVLLQHPNIFDAGVAGLPPDDAG 480
DB 421 HTGDIGYYDEEKHFFIVDRILKSLIKYGYOVPPALESVLLQHPNIFDAGVAGLPPDDAG 480

Db 421 HTGDIGYYDEEKHFFIVDRILKSLIKYGYOVPPALESVLLQHPNIFDAGVAGLPPDDAG 480
QY 481 ELPGAAYVYLKKGKSMTEKEVMDIVASOVSNAKRLRGVRFVDEVPGLGKIDGKAIREI 540
Db 481 ELPGAAYVYLKKGKSMTEKEVMDIVASOVSNAKRLRGVRFVDEVPGLGKIDGKAIREI 540
QY 541 LKK 543
Db 541 LKK 543

Search completed: November 29, 2001, 01:53:47
Job time: 3207 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:53:47 ; Search time 26.42 Seconds
(without alignments)
466.760 Million cell updates/sec

Title: US-09-581-241-6

Perfect score: 2823
Sequence: 1 MENMENDNIYGEPEFYPI.....TGKIDGKAIKREILKKPVAKM 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2817	99.8	548	2 US-08-460-934-2	Sequence 2, Appl
2	2817	99.8	548	2 US-08-782-118-2	Sequence 2, Appl
3	2815	99.7	548	3 US-09-111-752-14	Sequence 14, Appl
4	2814	99.7	548	4 US-07-903-047-8	Sequence 8, Appl
5	2814	99.7	548	4 US-09-380-061B-16	Sequence 16, Appl
6	2812	99.6	636	2 US-08-460-934-9	Sequence 9, Appl
7	2812	99.6	636	2 US-08-782-118-9	Sequence 9, Appl
8	2799	99.1	568	2 US-08-460-934-6	Sequence 6, Appl
9	2799	99.1	568	2 US-08-782-118-6	Sequence 6, Appl
10	2701	95.7	552	3 US-09-111-752-10	Sequence 10, Appl
11	2684	95.1	548	1 US-07-675-211-2	Sequence 2, Appl
12	2684	95.1	548	1 US-07-903-047-2	Sequence 2, Appl
13	2684	95.1	548	1 US-08-076-042-2	Sequence 2, Appl
14	2684	95.1	548	1 US-09-380-061B-14	Sequence 14, Appl
15	2590	91.7	552	3 US-09-111-752-7	Sequence 7, Appl
16	2394	84.8	548	4 US-09-111-752-5	Sequence 5, Appl
17	2335	82.7	548	4 US-09-380-061B-18	Sequence 18, Appl
18	1967	69.7	547	4 US-09-380-061B-20	Sequence 20, Appl
19	1956	69.3	550	1 US-08-354-240A-4	Sequence 4, Appl
20	1947	69.0	550	1 US-08-354-240A-6	Sequence 6, Appl
21	1945	68.9	549	1 US-08-354-240A-2	Sequence 2, Appl
22	1945	68.9	550	3 US-08-867-352-23	Sequence 23, Appl
23	1945	68.9	550	4 US-09-380-061B-21	Sequence 21, Appl
24	1945	68.9	815	1 US-08-122-550C-9	Sequence 9, Appl
25	1942	68.8	550	4 US-08-718-425-5	Sequence 5, Appl
26	1940	68.7	550	4 US-08-875-277A-4	Sequence 4, Appl
27	1939	68.7	550	4 US-08-718-425-2	Sequence 2, Appl

28	1939.5	68.7	550	4 US-08-875-277A-2	Sequence 2, Appl
29	1937.5	68.6	550	4 US-09-380-061B-6	Sequence 6, Appl
30	1905.5	67.5	561	2 US-08-474-169-8	Sequence 8, Appl
31	1831	64.9	552	1 US-08-231-729B-6	Sequence 6, Appl
32	723.5	25.6	540	4 US-08-991-677-8	Sequence 8, Appl
33	587	20.8	578	4 US-08-981-215-1	Sequence 1, Appl
34	260.5	9.2	798	2 US-08-222-617A-8	Sequence 8, Appl
35	256	9.1	582	3 US-08-403-852D-20	Sequence 20, Appl
36	256	9.1	582	3 US-08-510-646B-21	Sequence 20, Appl
37	256	9.1	582	3 US-09-231-818-20	Sequence 20, Appl
38	251.5	8.9	669	2 US-07-861-800-2	Sequence 2, Appl
39	251	8.9	2618	4 US-09-413-814-28	Sequence 28, Appl
40	249	8.8	2628	4 US-09-413-814-11	Sequence 11, Appl
41	245	8.7	3665	2 US-08-222-617A-13	Sequence 13, Appl
42	245	8.7	3712	2 US-08-222-617A-4	Sequence 4, Appl
43	245	8.7	3712	2 US-08-222-617A-25	Sequence 25, Appl
44	233	8.3	768	2 US-08-222-617A-5	Sequence 5, Appl
45	233	8.3	3666	2 US-08-222-617A-12	Sequence 12, Appl

ALIGNMENTS

```

RESULT
1
US-08-460-934-2
; Sequence 2, Application US/08460934
; Patent No. 5814465
GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: KOTAMA, YASUJI
TITLE OF INVENTION: BIOTINATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINATED AND A BIOLUMINESCENT
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESS: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,934
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids

```

```

: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: peptide
: ORIGINAL SOURCE:
: ORGANISM: Luciola lateralis
US-08-460-934-2

```

```

Query Match          99.8%; Score 2817; DB 2; Length 548;
Best local similarity 99.6%; Pred. No. 6.6e-286;
Matches 546; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MENMENDENIVYGEPEPEPIEESGAGOLRKRYMDRAKGAIAFTNALTVDTYAYEYLE 60
DB 1 MENMENDENIVYGEPEPEPIEESGAGOLRKRYMDRAKGAIAFTNALTVDTYAYEYLE 60
QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPLVAGLFIYGAVAPNEIYTLRELHVS 120
DB 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPLVAGLFIYGAVAPNEIYTLRELHVS 120
QY 121 GISKPTIVSSSKGDLKVTVOKTVAIKTIYILDSKVDYRGYOSMDNFKKNTPOGFKG 180
DB 121 GISKPTIVSSSKGDLKVTVOKTVAIKTIYILDSKVDYRGYOSMDNFKKNTPOGFKG 180
QY 181 SSKTYEVNRKEOVALLIMSSGSTGLPKGVOLTHENIVTRFSHARDPIYGNVSPGTAIL 240
DB 181 SSKTYEVNRKEOVALLIMSSGSTGLPKGVOLTHENIVTRFSHARDPIYGNVSPGTAIL 240
QY 241 TVPFFHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTDYKCSSVILVPTLAILNRS 300
DB 241 TVPFFHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTDYKCSSVILVPTLAILNRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRNPLPGVROGYGLTETTSAILITPEGDDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRNPLPGVROGYGLTETTSAILITPEGDDK 360
QY 361 PASGSKVVPFLFAKVIDDTTKTLGPNRRGEVCVGPMLMKGYVNDPEATREIIDEEGWL 420
DB 361 PASGSKVVPFLFAKVIDDTTKTLGPNRRGEVCVGPMLMKGYVNDPEATREIIDEEGWL 420
QY 421 HNGDIGYUDEKHEFTVDRLSLTKYKGYOVPAELDESILLHPNIFPDAGVAPDPDIA 480
DB 421 HNGDIGYUDEKHEFTVDRLSLTKYKGYOVPAELDESILLHPNIFPDAGVAPDPDIA 480
QY 481 ELPGAVVYLKKGKSMTEKEVMDYVASOVNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
DB 481 ELPGAVVYLKKGKSMTEKEVMDYVASOVNAKRLRGVRFVDEVPKGLTGKIDGKAIREI 540
QY 541 LKKPVAKM 548
DB 541 LKKPVAKM 548

```

```

RESULT 2
US-08-782-118-2
: Sequence 2, Application US/08782118
: Patent No. 5843746

```

```

: GENERAL INFORMATION:
: APPLICANT: TATSUMI, HIROKI
: APPLICANT: FUKUDA, SATOSHI
: APPLICANT: KIKUCHI, MAMORU
: APPLICANT: KOYAMA, YASUJI
: TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
: TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
: TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESSES:
: ADDRESS: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
: ADDRESS: P.C. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
: STREET: 1755 S.
: CITY: ARLINGTON

```

```

: STATE: VA
: COUNTRY: USA
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/782,118
: FILING DATE: 13-JAN-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/460,934
: FILING DATE: 05-JUN-1995
: APPLICATION NUMBER: JP 193798/1994
: FILING DATE: 27-JUL-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 54625/1995
: FILING DATE: 14-MAR-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 98857/1995
: FILING DATE: 24-APR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: OBLON, NORMAN F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 7126-001-0
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-413-3000
: TELEFAX: 703-413-2220
: INFORMATION FOR SEO ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 548 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: MOLECULE TYPE: unknown
: ORIGINAL SOURCE:
: ORGANISM: Luciola lateralis
US-08-782-118-2

```

```

Query Match          99.8%; Score 2817; DB 2; Length 548;
Best local similarity 99.6%; Pred. No. 6.6e-286;
Matches 546; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MENMENDENIVYGEPEPEPIEESGAGOLRKRYMDRAKGAIAFTNALTVDTYAYEYLE 60
DB 1 MENMENDENIVYGEPEPEPIEESGAGOLRKRYMDRAKGAIAFTNALTVDTYAYEYLE 60
QY 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPLVAGLFIYGAVAPNEIYTLRELHVS 120
DB 61 KSCCLGEALKNYGLVVDGRIALCSENCEEFFIPLVAGLFIYGAVAPNEIYTLRELHVS 120
QY 121 GISKPTIVSSSKGDLKVTVOKTVAIKTIYILDSKVDYRGYOSMDNFKKNTPOGFKG 180
DB 121 GISKPTIVSSSKGDLKVTVOKTVAIKTIYILDSKVDYRGYOSMDNFKKNTPOGFKG 180
QY 181 SSKTYEVNRKEOVALLIMSSGSTGLPKGVOLTHENIVTRFSHARDPIYGNVSPGTAIL 240
DB 181 SSKTYEVNRKEOVALLIMSSGSTGLPKGVOLTHENIVTRFSHARDPIYGNVSPGTAIL 240
QY 241 TVPFFHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTDYKCSSVILVPTLAILNRS 300
DB 241 TVPFFHGFMTTLLGYLTCGFRIVMLTKFDEETFLKTDYKCSSVILVPTLAILNRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRNPLPGVROGYGLTETTSAILITPEGDDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARRNPLPGVROGYGLTETTSAILITPEGDDK 360
QY 361 PASGSKVVPFLFAKVIDDTTKTLGPNRRGEVCVGPMLMKGYVNDPEATREIIDEEGWL 420
DB 361 PASGSKVVPFLFAKVIDDTTKTLGPNRRGEVCVGPMLMKGYVNDPEATREIIDEEGWL 420

```


QY 421 HTGDIGYDEEKKHFTVDRLKSLIKYQVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 |||||
 DB 421 HTGDIGYDEEKKHFTVDRLKSLIKYQVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 QY 481 ELPGAVVVLKKGKSMTEKEVMDYVASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIRFI 540
 |||||
 DB 481 ELPGAVVVLKKGKSMTEKEVMDYVASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIRFI 540
 QY 541 LKRPVAKM 548
 |||||
 DB 541 LKRPVAKM 548
 RESULT 3
 US-09-111-752-14
 ; Sequence 14, Application US/09111752
 ; Patent No. 6074859
 ; GENERAL INFORMATION:
 ; APPLICANT: HIROKAWA, KOZO
 ; APPLICANT: KAJIYAMA, NAOKI
 ; APPLICANT: MURAKAMI, SEIJI
 ; TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
 ; TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE LUMINESCENT PROTEIN
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ; ADDRESSEE: P.C.
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
 ; CITY: ARLINGTON
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/111,752
 ; FILING DATE: 08-JUL-1998
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F.
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 7126-0009-0
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-413-3000
 ; TELEFAX: 703-413-2220
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 548 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-111-752-14
 Query Match 99.7%; Score 2815; DB 3; Length 548;
 Best Local Similarity 99.6%; Pred. No. 1.1e-285;
 Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 121 GISKPTIVSSKKGLDKVTYVQKTAITKTIYILDSKYDVRGYSMDNFIKKNTPOGFKG 180
 QY 181 SSKPTIVEVNRKQVALLINSSGSTGLPKGVOLTHENTVTRSHADPTLYGNQVSPGTAIL 240
 |||||
 DB 181 SSKPTIVEVNRKQVALLINSSGSTGLPKGVOLTHENTVTRSHADPTLYGNQVSPGTAIL 240
 QY 241 TVVPRFHGFMFTTGLTGLCGFRIMLTKFDEEFTLKTLDYKCSSVLLVPTLFAILNRS 300
 |||||
 DB 241 TVVPRFHGFMFTTGLTGLCGFRIMLTKFDEEFTLKTLDYKCSSVLLVPTLFAILNRS 300
 QY 301 ELLDKYDLSNLEIASGAPLSKEIGEAVARRNLPVGRVGGTLETTSAIITPEGDGK 360
 |||||
 DB 301 ELLDKYDLSNLEIASGAPLSKEIGEAVARRNLPVGRVGGTLETTSAIITPEGDGK 360
 QY 361 PGASGVVPLFPAKYIDLDTKTLPNRRGEVGVGPHLMKYVNNPNTREITIDEEGML 420
 |||||
 DB 361 PGASGVVPLFPAKYIDLDTKTLPNRRGEVGVGPHLMKYVNNPNTREITIDEEGML 420
 QY 421 HTGDIGYDEEKKHFTVDRLKSLIKYQVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 |||||
 DB 421 HTGDIGYDEEKKHFTVDRLKSLIKYQVPPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 QY 481 ELPGAVVVLKKGKSMTEKEVMDYVASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIRFI 540
 |||||
 DB 481 ELPGAVVVLKKGKSMTEKEVMDYVASQVSNARKRLRGVRFVDEVPKGLTGKIDGKAIRFI 540
 QY 541 LKRPVAKM 548
 |||||
 DB 541 LKRPVAKM 548
 RESULT 4
 US-07-903-047-8
 ; Sequence 8, Application US/07903047
 ; Patent No. 5229285
 ; GENERAL INFORMATION:
 ; APPLICANT: Kajiya, Naoki
 ; APPLICANT: Nakano, Eiichi
 ; TITLE OF INVENTION: Thermostable Luciferase of Firefly,
 ; TITLE OF INVENTION: DNA, and Process for The Preparation of Thermostable
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/903,047
 ; FILING DATE: 19920623
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mastrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 7005-048
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-9090
 ; TELEFAX: 212 869-8864/9741
 ; TELEEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 548 amino acids
 ; TYPE: AMINO ACID
 ; STRANDEDNESS: single
 ; TOPOLOGY: unknown

MOLECULE TYPE: peptide
US-07-903-047-8

Query Match
Best Local Similarity 99.7%; Score 2814; DB 1; Length 548;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MENMENDENIVYGPPEFYPIEESGAGOLRKYMRYAKIGALIAFTNALGVDTVAEYLE 60
DB 1 MEMMENDENIVYGPPEFYPIEESGAGOLRKYMRYAKIGALIAFTNALGVDTVAEYLE 60
QY 61 KSCCGLGALKNGLVVDGRIALCSENCEEFFIPVLAGLFGVAVPTNEIYTLRELIVHSL 120
DB 61 KSCCGLGALKNGLVVDGRIALCSENCEEFFIPVLAGLFGVAVPTNEIYTLRELIVHSL 120
QY 121 GISKPTIVSSKKGLDVTYOKTVAIKTIYILDSKVYRGQSDMNFIKNTPOGFGK 180
DB 121 GISKPTIVSSKKGLDVTYOKTVAIKTIYILDSKVYRGQSDMNFIKNTPOGFGK 180
QY 181 SSFKTVEVNRKQVALINMSSGSTGLPKGVOLTHENIYTRFSHARDPIYGNVSPGTAIL 240
DB 181 SSFKTVEVNRKQVALINMSSGSTGLPKGVOLTHENIYTRFSHARDPIYGNVSPGTAIL 240
QY 241 TVPFRHGGFMFTTGLYLCGFRIVMLTKFDEETPLKTLQDYKCSSVILVPTLFAILNRS 300
DB 241 TVPFRHGGFMFTTGLYLCGFRIVMLTKFDEETPLKTLQDYKCSSVILVPTLFAILNRS 300
QY 301 ELLDKYDLSNLEIASGAPLSKEIGEAVARRPNLPGVROGVLTEETSAIITPEGDDK 360
DB 301 ELLDKYDLSNLEIASGAPLSKEIGEAVARRPNLPGVROGVLTEETSAIITPEGDDK 360
QY 361 PGASGVVPLFRKAKYIDDTKTKTLPNRRGEVCVGPMLMGVYDNPATREIIDEBSWL 420
DB 361 PGASGVVPLFRKAKYIDDTKTKTLPNRRGEVCVGPMLMGVYDNPATREIIDEBSWL 420
QY 421 HTGDIGYDEEKNHFTVDRKSLIKYKGYVPALESVLLOHPNIFDAGVAVDPPIAG 480
DB 421 HTGDIGYDEEKNHFTVDRKSLIKYKGYVPALESVLLOHPNIFDAGVAVDPPIAG 480
QY 481 ELPGAIVVLLKKGKSMTEKEVMDYVASOVSNARKLRGVRVDEVKGLTKIDGKAIREI 540
DB 481 ELPGAIVVLLKKGKSMTEKEVMDYVASOVSNARKLRGVRVDEVKGLTKIDGKAIREI 540
QY 541 LKRPVAKM 548
DB 541 LKRPVAKM 548

```

RESULT 5
US-09-380-061B-16
Sequence 16, Application US/09380061B
Patent No. 6265177

GENERAL INFORMATION:

APPLICANT: SOUTIRRELL, DAVID JAMES
WHITE, PETER JOHN
LOWE, CHRISTOPHER ROBIN
MURRAY, JAMES AUGUSTUS HENRY
TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/380,061B
FILING DATE: 25-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB98/01026
FILING DATE: 7-APR-1998
APPLICATION NUMBER: GB 9707468.8
FILING DATE: 11-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B. J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 124-725
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4000
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-380-061B-16

Query Match
Best Local Similarity 99.7%; Score 2814; DB 4; Length 548;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MENMENDENIVYGPPEFYPIEESGAGOLRKYMRYAKIGALIAFTNALGVDTVAEYLE 60
DB 1 MEMMENDENIVYGPPEFYPIEESGAGOLRKYMRYAKIGALIAFTNALGVDTVAEYLE 60
QY 61 KSCCGLGALKNGLVVDGRIALCSENCEEFFIPVLAGLFGVAVPTNEIYTLRELIVHSL 120
DB 61 KSCCGLGALKNGLVVDGRIALCSENCEEFFIPVLAGLFGVAVPTNEIYTLRELIVHSL 120
QY 121 GISKPTIVSSKKGLDVTYOKTVAIKTIYILDSKVYRGQSDMNFIKNTPOGFGK 180
DB 121 GISKPTIVSSKKGLDVTYOKTVAIKTIYILDSKVYRGQSDMNFIKNTPOGFGK 180
QY 181 SSFKTVEVNRKQVALINMSSGSTGLPKGVOLTHENIYTRFSHARDPIYGNVSPGTAIL 240
DB 181 SSFKTVEVNRKQVALINMSSGSTGLPKGVOLTHENIYTRFSHARDPIYGNVSPGTAIL 240
QY 241 TVPFRHGGFMFTTGLYLCGFRIVMLTKFDEETPLKTLQDYKCSSVILVPTLFAILNRS 300
DB 241 TVPFRHGGFMFTTGLYLCGFRIVMLTKFDEETPLKTLQDYKCSSVILVPTLFAILNRS 300
QY 301 ELLDKYDLSNLEIASGAPLSKEIGEAVARRPNLPGVROGVLTEETSAIITPEGDDK 360
DB 301 ELLDKYDLSNLEIASGAPLSKEIGEAVARRPNLPGVROGVLTEETSAIITPEGDDK 360
QY 361 PGASGVVPLFRKAKYIDDTKTKTLPNRRGEVCVGPMLMGVYDNPATREIIDEBSWL 420
DB 361 PGASGVVPLFRKAKYIDDTKTKTLPNRRGEVCVGPMLMGVYDNPATREIIDEBSWL 420
QY 421 HTGDIGYDEEKNHFTVDRKSLIKYKGYVPALESVLLOHPNIFDAGVAVDPPIAG 480
DB 421 HTGDIGYDEEKNHFTVDRKSLIKYKGYVPALESVLLOHPNIFDAGVAVDPPIAG 480
QY 481 ELPGAIVVLLKKGKSMTEKEVMDYVASOVSNARKLRGVRVDEVKGLTKIDGKAIREI 540
DB 481 ELPGAIVVLLKKGKSMTEKEVMDYVASOVSNARKLRGVRVDEVKGLTKIDGKAIREI 540
QY 541 LKRPVAKM 548
DB 541 LKRPVAKM 548

```

RESULT 6
US-08-460-934-9
Sequence 9, Application US/08460934

```

Patent No. 5814465
GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: KOYAMA, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
TITLE OF INVENTION: ANALYSIS METHOD
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,934
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-934-9

```

```

Query Match          99.6%; Score 2812; DB 2; Length 636;
Best Local Similarity 99.6%; Pred. No. 2.9e-285;
Matches 545; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MENMENDENIYVGPPEFYPIEGSGAGQIRKYMDRYAKGATAFNATLGVDYTAETLE 60
DB 1 MENMENDENIYVGPPEFYPIEGSGAGQIRKYMDRYAKGATAFNATLGVDYTAETLE 60
QY 61 KSCGCEALKNYGLVVDGRIALSCENCEFFIPVLAGLFTIGVVAFTNIEYTLRELHSL 120
DB 61 KSCGCEALKNYGLVVDGRIALSCENCEFFIPVLAGLFTIGVVAFTNIEYTLRELHSL 120
QY 121 GISKPTIVSSKKGDKLVTVQKTYATITVILDSKVDYRGYQSMDFIKNTTQGFPG 180
DB 121 GISKPTIVSSKKGDKLVTVQKTYATITVILDSKVDYRGYQSMDFIKNTTQGFPG 180
QY 181 SSFKTYEVNRKEOVALLIMSSGSTGLPKGVOLTHEMLVTRFSHARDPITGNVSPGTALL 240
DB 181 SSFKTYEVNRKEOVALLIMSSGSTGLPKGVOLTHEMLVTRFSHARDPITGNVSPGTALL 240

```

```

QY 241 TVPFFHGFMTFTGLYLTCGFRIVMLTKFDEETFLKTLQDYKCSSVLLVPTLAILNRS 300
DB 241 TVPFFHGFMTFTGLYLTCGFRIVMLTKFDEETFLKTLQDYKCSSVLLVPTLAILNRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARFNPYGRQCYGLTETTSAILITPEGDG 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARFNPYGRQCYGLTETTSAILITPEGDG 360
QY 361 PGASGVVPLFAKAVYDLDTKTKTGPNRGECYVGPMLMKGYVNPATREIIDEGBGL 420
DB 361 PGASGVVPLFAKAVYDLDTKTKTGPNRGECYVGPMLMKGYVNPATREIIDEGBGL 420
QY 421 HTGDIGYDEKHFETVDRLSLTKYGYVPPAELESYLLQHPNIFDAGVAGVDPDIAG 480
DB 421 HTGDIGYDEKHFETVDRLSLTKYGYVPPAELESYLLQHPNIFDAGVAGVDPDIAG 480
QY 481 ELPGAVVYLKKGKSMTEKEVMYVAVSNAKRIKGVAFVDEYKGLTKIDGKAIREI 540
DB 481 ELPGAVVYLKKGKSMTEKEVMYVAVSNAKRIKGVAFVDEYKGLTKIDGKAIREI 540
QY 541 LKKPVAK 547
DB 541 LKKPVAK 547

```

```

RESULT 7
US-08-782-118-9
Sequence 9, Application US/08782118
Patent No. 5843746
GENERAL INFORMATION:
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: KOYAMA, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/782,118
FILING DATE: 13-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,934
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000

```

TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 636 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-782-118-9

Query Match 99.6%; Score 2812; DB 2; Length 636;
 Best Local Similarity 99.6%; Pred. No. 2.9e-285;
 Matches 545; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MENMENDENIVGPPEFYPIEESAGAOIRKYMRYAKIGALFTNALGVDTYAEYLE 60
DB 1 MENMENDENIVGPPEFYPIEESAGAOIRKYMRYAKIGALFTNALGVDTYAEYLE 60
QY 61 KSCCLGEBALKNGLVVDGRALALSCENCEFFIPVLAGLFIGVAPFNEITLRELHSL 120
DB 61 KSCCLGEBALKNGLVVDGRALALSCENCEFFIPVLAGLFIGVAPFNEITLRELHSL 120
QY 121 GISKPPIVSSKKGLDKVITVOKTVAITVITLDSKVYRGYQSDNFIKNTPOGFGK 180
DB 121 GISKPPIVSSKKGLDKVITVOKTVAITVITLDSKVYRGYQSDNFIKNTPOGFGK 180
QY 181 SSKFYEVNRRKEQVALINMSSSTGLPKGVOLTHENITVRFSHARDPIYGNVSPGTAIL 240
DB 181 SSKFYEVNRRKEQVALINMSSSTGLPKGVOLTHENITVRFSHARDPIYGNVSPGTAIL 240
QY 241 TVVPHFGMFTTGLYLCGFRIVMLTFDEDEFLKTIQDYKSSVILVPLFALINLS 300
DB 241 TVVPHFGMFTTGLYLCGFRIVMLTFDEDEFLKTIQDYKSSVILVPLFALINLS 300
QY 301 ELLDKYDLSNLEIASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDGK 360
DB 301 ELLDKYDLSNLEIASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDGK 360
QY 361 PASGKVVNPLFAKVIDLDTKTLGPNRGEVCKGPMIMKGYVNDPEATREIIEEGWL 420
DB 361 PASGKVVNPLFAKVIDLDTKTLGPNRGEVCKGPMIMKGYVNDPEATREIIEEGWL 420
QY 421 HNGDIGYDEEHFFIVDLKSLIKYKGYVPALEESVILQHPNIFPAGVAGVDPDTAG 480
DB 421 HNGDIGYDEEHFFIVDLKSLIKYKGYVPALEESVILQHPNIFPAGVAGVDPDTAG 480
QY 481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNAKRLGVRVDEVPKGLTGKIDGKAIREI 540
DB 481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNAKRLGVRVDEVPKGLTGKIDGKAIREI 540
QY 541 LKPPYAK 547
DB 541 LKPPYAK 547

```

RESULT 8
 US-08-460-934-6
 Sequence 6, Application US/08460934
 Patent No. 5814465

GENERAL INFORMATION:
 APPLICANT: TATSUMI, HIROKI
 APPLICANT: FUKUDA, SATOSHI
 APPLICANT: KIRUCHI, MAMORU
 APPLICANT: KOYAMA, YASUJI
 TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
 TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
 TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR

CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patentln Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/460,934
 FILING DATE: 05-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 193798/1994
 FILING DATE: 27-JUL-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 54625/1995
 FILING DATE: 14-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 98857/1995
 FILING DATE: 24-APR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 7126-001-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 568 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-460-934-6

Query Match 99.1%; Score 2799; DB 2; Length 568;
 Best Local Similarity 99.3%; Pred. No. 5.4e-284;
 Matches 542; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 MNENDENIVGPPEFYPIEESAGAOIRKYMRYAKIGALFTNALGVDTYAEYLE 62
DB 3 MNENDENIVGPPEFYPIEESAGAOIRKYMRYAKIGALFTNALGVDTYAEYLE 62
QY 63 CCLGEBALKNGLVVDGRALALSCENCEFFIPVLAGLFIGVAPFNEITLRELHSLGI 122
DB 63 CCLGEBALKNGLVVDGRALALSCENCEFFIPVLAGLFIGVAPFNEITLRELHSLGI 122
QY 123 SKPTIVSSKKGLDKVITVOKTVAITVITLDSKVYRGYQSDNFIKNTPOGFGKSS 182
DB 123 SKPTIVSSKKGLDKVITVOKTVAITVITLDSKVYRGYQSDNFIKNTPOGFGKSS 182
QY 183 SKPTIVSSKKGLDKVITVOKTVAITVITLDSKVYRGYQSDNFIKNTPOGFGKSS 202
DB 183 SKPTIVSSKKGLDKVITVOKTVAITVITLDSKVYRGYQSDNFIKNTPOGFGKSS 202
QY 203 FKTVEVNRKEQVALINMSSSTGLPKGVOLTHENITVRFSHARDPIYGNVSPGTAILTV 262
DB 203 FKTVEVNRKEQVALINMSSSTGLPKGVOLTHENITVRFSHARDPIYGNVSPGTAILTV 262
QY 243 VPHFGMFTTGLYLCGFRIVMLTFDEDEFLKTIQDYKSSVILVPLFALINRSL 302
DB 243 VPHFGMFTTGLYLCGFRIVMLTFDEDEFLKTIQDYKSSVILVPLFALINRSL 302
QY 303 LDKYDLSNLEIASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDGK 362
DB 303 LDKYDLSNLEIASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDGK 362
QY 323 LDKYDLSNLEIASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDGK 382
DB 323 LDKYDLSNLEIASGAPLSKEIGEAVARFNLPGVROGGLTETTSAILITPEGDGK 382
QY 363 ASKVPLFAKVIDLDTKTLGPNRGEVCKGPMIMKGYVNDPEATREIIEEGWLHT 422
DB 363 ASKVPLFAKVIDLDTKTLGPNRGEVCKGPMIMKGYVNDPEATREIIEEGWLHT 422
QY 423 GDIGYDEEHFFIVDLKSLIKYKGYVPALEESVILQHPNIFPAGVAGVDPDTAGEL 482
DB 423 GDIGYDEEHFFIVDLKSLIKYKGYVPALEESVILQHPNIFPAGVAGVDPDTAGEL 482
QY 443 GDIGYDEEHFFIVDLKSLIKYKGYVPALEESVILQHPNIFPAGVAGVDPDTAGEL 502
DB 443 GDIGYDEEHFFIVDLKSLIKYKGYVPALEESVILQHPNIFPAGVAGVDPDTAGEL 502

```

Qy	Db	Qy	Db
483	503	543	563
PGAAVVVLKKGKSMTEK	PGAAVVVLEKGSMTKE	KPVAKM	KPVAKM
EVMDYVASQVSNARKL	EMDYVASQVSNARKL	IIIIII	IIIIII
RCGVRFVDEVPGLGKID	RCGVRFVDEVPGLGKID	IIIIII	IIIIII
KAIREILK	KAIREILK		

RESULT 9
US-08-782-118-6
; Sequence 6, Application US/08782118
; Patent No. 5843746

QUANTITY: 1000000
APPLICANT: TATSUMI, HIROKI
APPLICANT: FUKUDA, SATOSHI
APPLICANT: KIKUCHI, MAMORU
APPLICANT: Koyama, YASUJI
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA,
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINATED AND A BIOLUMINESCENT
TITLE OF INVENTION: ANALYSIS METHOD
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/782,118
FILING DATE: 13-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,934
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 568 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-782-118-6

Query Match	99.1%;	Score 2799;	DB 2;	Length 568;
Best Local Similarity	99.3%;	Pred. No. 5,4e-204;		
Matches 542;	Conservative	4;	Mismatches 0;	Indels 0;
Gaps 0;				

Db	23	SLDENDEIVIVGPPRPYIEEGSAGADLRKMDRYALKAIAFTNALTGVDYVYAEYLEKS	82
QY	63	CCUGALKNKGLVVDGRIALCSENCBEEFIPVLAGFIGVGAPNTEITRLRLVSLGI	122
Db	83	CCUGALKNKGLVVDGRIALCSENCBEEFIPVLAGFIGVGAPNTEITRLRLVSLGI	142
QY	123	SKPIIVSSKKGDKVTTYOKTVAIKTIVILDSKYDRGQSMDFIKKNTPOGFKGSS	182
Db	143	SKPIIVSSKKGDKVTTYOKTVAIKTIVILDSKYDRGQSMDFIKKNTPOGFKGSS	202
QY	183	EKTYEVNRKEQVALINMSSGSTGLPKGVOLTHENITVRFSHADPIYGNVSPGTALIV	242
Db	203	EKTYEVNRKEQVALINMSSGSTGLPKGVOLTHENITVRFSHADPIYGNVSPGTALIV	262
QY	243	VPIFHGGMFTTLGTYLTCGPRIVMLTGFDEDETFKLTDYDKCSSVILVPLFLINRSEL	302
Db	263	VPIFHGGMFTTLGTYLTCGPRIVMLTGFDEDETFKLTDYDKCSSVILVPLFLINRSEL	322
QY	303	LDKXDSLNLVEIASGAPALSKETIGEAVARRENPGVBOGYGLTETTSAILITPEBGDKPG	362
Db	323	LDKXDSLNLVEIASGAPALSKETIGEAVARRENPGVBOGYGLTETTSAILITPEBGDKPG	382
QY	363	ASGVAVPLFAKKAYIDLDTKTLGPNRKGVEYVKKPMLMKGVYNDPEATREIIEEGMLHT	422
Db	383	ASGVAVPLFAKKAYIDLDTKTLGPNRKGVEYVKKPMLMKGVYNDPEATREIIEEGMLHT	442
QY	423	GDIGYDEEKEHFEIYDRLSLITKKGYVPPAELESVLLQHPNFTPDGAVGVPDPIAGEL	482
Db	443	GDIGYDEEKEHFEIYDRLSLITKKGYVPPAELESVLLQHPNFTPDGAVGVPDPIAGEL	502
QY	483	PGAVVVLKKGSMTKEEYMDIVYASOVSAKRLRGGVAVDEVPPGGLGKIDGKAIARELIK	542
Db	503	PGAVVVLKKGSMTKEEYMDIVYASOVSAKRLRGGVAVDEVPPGGLGKIDGKAIARELIK	562
QY	543	KPVAKM	548
Db	563	KPVAKM	568

RESULT 10
 US-09-111-752-10
 ; Sequence 10, Application US/09111752
 ; Patent No. 6074859
 ; GENERAL INFORMATION:
 APPLICANT: HIROKAWA, KOZO
 APPLICANT: KAJIYAMA, NAOKI
 APPLICANT: MORAKAMI, SEIJI
 TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
 TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE LUMINESCENT PROTEIN
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
 CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/111,752
 FILING DATE: 08-JUL-1998
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 7126-0009-0
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 552 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola lateralis, Photinus pyralis
 US-09-111-752-10

Query Match 95.7%; Score 2701; DB 3; Length 552;
 Best Local Similarity 96.1%; Pred. No. 9.3e-274;
 Matches 522; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

```

QY 1 MNNMNDENIVGPEPPYPIEBSAGAOQLRKYMDRAKAGALFTNALGVDTYAETYLE 60
DB 1 MNNMNDENIVGPEPPYPIEBSAGAOQLRKYMDRAKAGALFTNALGVDTYAETYLE 60
QY 61 KSCCLGALKNYGLVVDGRIALCSENCCEFFIPVLAGLFIGVAPNTNITLRELHSL 120
DB 61 KSCCLGALKNYGLVVDGRIALCSENCCEFFIPVLAGLFIGVAPNTNITLRELHSL 120
QY 121 GISKPTIVSSKKGLDKVTVOKTVAITKIYILDSKVYRGYSMDNFIKKNTPQGFKG 180
DB 121 GISKPTIVSSKKGLDKVTVOKTVAITKIYILDSKVYRGYSMDNFIKKNTPQGFKG 180
QY 181 SSFKTEVNRKKEQVALINMSGSTGLPKGVOLTHENIVRFSHARDPIYGNVSGTAL 240
DB 181 SSFKTEVNRKKEQVALINMSGSTGLPKGVOLTHENIVRFSHARDPIYGNVSGTAL 240
QY 241 TVVPHHGFMTTLGYLTGCFRIYMLTKFDEETPLKTLQDYKCSSVLVPLFAILNRS 300
DB 241 TVVPHHGFMTTLGYLTGCFRIYMLTKFDEETPLKTLQDYKCSSVLVPLFAILNRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARPNLPVROGYGLTETTSAILITPBGDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARPNLPVROGYGLTETTSAILITPBGDK 360
QY 361 PGASGVVPLFRAKVIDDTKTKLGPNNRGEVCVGPMLMGYVNDNPNATREIIDEEGWL 420
DB 361 PGASGVVPLFRAKVIDDTKTKLGPNNRGEVCVGPMLMGYVNDNPNATREIIDEEGWL 420
QY 421 HTGDIGYDEEKHFIVRLKSLIKYKGYVPAPLESVLLQHPNIDAGVAGVDDPTAG 480
DB 421 HTGDIGYDEEKHFIVRLKSLIKYKGYVPAPLESVLLQHPNIDAGVAGVDDPTAG 480
QY 481 ELPGAIVVLKKGKSKTEKEVMDYVASQVSNKRLRGVRFVDEVPKGLTGKIDKAIKREI 540
DB 481 ELPGAIVVLKKGKSKTEKEVMDYVASQVSNKRLRGVRFVDEVPKGLTGKIDKAIKREI 540
QY 541 LKK 543
DB 541 LKK 543

```

RESULT 11
 US-07-675-211-2
 Sequence 2, Application US/07675211
 Patent No. 5219737
 GENERAL INFORMATION:
 APPLICANT: KAJIYAMA, NAOKI
 TITLE OF INVENTION: MUTANT LUCIFERASE OF A FIREFLY, MUTANT
 TITLE OF INVENTION: LUCIFERASE GENES, NOVEL RECOMBINANT DNAS CONTAINING THE
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PENNIE & EDMONDS
 STREET: 1155 AVENUE OF THE AMERICAS

CITY: NEW YORK
 STATE: N.Y.
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/675,211
 FILING DATE: 19910326
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MISROCK, S. LESLIE
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7005-026-999
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 548 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola cruciata
 US-07-675-211-2

Query Match 95.1%; Score 2684; DB 1; Length 548;
 Best Local Similarity 93.4%; Pred. No. 5.5e-272;
 Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

```

QY 1 MNNMNDENIVGPEPPYPIEBSAGAOQLRKYMDRAKAGALFTNALGVDTYAETYLE 60
DB 1 MNNMNDENIVGPKPPYPIEBSAGAOQLRKYMERAKAGALFTNAVGVDTYAETYLE 60
QY 61 KSCCLGALKNYGLVVDGRIALCSENCCEFFIPVLAGLFIGVAPNTNITLRELHSL 120
DB 61 KSCCLGALKNYGLVVDGRIALCSENCCEFFIPVLAGLFIGVAPNTNITLRELHSL 120
QY 121 GISKPTIVSSKKGLDKVTVOKTVAITKIYILDSKVYRGYSMDNFIKKNTPQGFKG 180
DB 121 GISKPTIVSSKKGLDKVTVOKTVAITKIYILDSKVYRGYSMDNFIKKNTPQGFKG 180
QY 181 SSFKTEVNRKKEQVALINMSGSTGLPKGVOLTHENIVRFSHARDPIYGNVSGTAL 240
DB 181 SSFKTEVNRKKEQVALINMSGSTGLPKGVOLTHENIVRFSHARDPIYGNVSGTAL 240
QY 241 TVVPHHGFMTTLGYLTGCFRIYMLTKFDEETPLKTLQDYKCSSVLVPLFAILNRS 300
DB 241 TVVPHHGFMTTLGYLTGCFRIYMLTKFDEETPLKTLQDYKCSSVLVPLFAILNRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARPNLPVROGYGLTETTSAILITPBGDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARPNLPVROGYGLTETTSAILITPBGDK 360
QY 361 PGASGVVPLFRAKVIDDTKTKLGPNNRGEVCVGPMLMGYVNDNPNATREIIDEEGWL 420
DB 361 PGASGVVPLFRAKVIDDTKTKLGPNNRGEVCVGPMLMGYVNDNPNATREIIDEEGWL 420
QY 421 HTGDIGYDEEKHFIVRLKSLIKYKGYVPAPLESVLLQHPNIDAGVAGVDDPTAG 480
DB 421 HTGDIGYDEEKHFIVRLKSLIKYKGYVPAPLESVLLQHPNIDAGVAGVDDPTAG 480
QY 481 ELPGAIVVLKKGKSKTEKEVMDYVASQVSNKRLRGVRFVDEVPKGLTGKIDKAIKREI 540
DB 481 ELPGAIVVLKKGKSKTEKEVMDYVASQVSNKRLRGVRFVDEVPKGLTGKIDKAIKREI 540
QY 541 LKKPAAK 548

```

Db 541 LKRPVAKM 548

RESULT 12
US-07-903-047-2

; Sequence 2, Application US/07903047
; Patent No. 529285

; GENERAL INFORMATION:

; APPLICANT: Kajiya, Naoki

; APPLICANT: Nakano, Eiichi

; TITLE OF INVENTION: Thermostable Luciferase Of Firefly,

; TITLE OF INVENTION: DNA, and Process For The Preparation Of Thermostable

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/903,047

; FILING DATE: 19920623

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Mistrock, S. Leslie

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 7005-048

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212 790-9090

; TELEFAX: 212 869-8644/9741

; TELEFAX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 548 amino acids

; TYPE: AMINO ACID

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

; US-07-903-047-2

Query Match 95.1%; Score 2684; DB 1; Length 548;

Best Local Similarity 93.4%; Pred. No. 5.5e-272;

Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

QY 1 MENMENDEIVVGPPEPIIEGSGAGOLRKYMRYAKLGAIAFTNALGVDTYAEYLE 60
DB 1 MENMENDEIVVGPPEPIIEGSGAGOLRKYMRYAKLGAIAFTNALGVDTYAEYLE 60
QY 61 KSCCLGKALQNYGLVVDRIALCSNCEFFPIVLAGLFIGVAVPNEIYTLRLVHSL 120
DB 61 KSCCLGKALQNYGLVVDRIALCSNCEFFPIVLAGLFIGVAVPNEIYTLRLVHSL 120
QY 121 GISKPTIVSSKKGDKYITVQKTATKTYITLSDKYDRGOSMDNFKKNTPOGFKG 180
DB 121 GISKPTIVSSKKGDKYITVQKTATKTYITLSDKYDRGOSMDNFKKNTPOGFKG 180
QY 181 SSFKTYEVNRKROVALINSSSGTGLPRGVOLTHENIVTRFESHARDPIYGNVSGTAL 240
DB 181 SSFKTYEVNRKROVALINSSSGTGLPRGVOLTHENIVTRFESHARDPIYGNVSGTAL 240
QY 241 TVVPRHNGGATTGTLGCGFRIVMLTKFDEETFLKTLQDKSSVILVPTLAILMS 300
DB 241 TVVPRHNGGATTGTLGCGFRIVMLTKFDEETFLKTLQDKSSVILVPTLAILMS 300

QY 301 ELLDKYDLSNIVEIASGAPLSKEIGEAVARRENLPGRVGGLTETTSATITPEGDK 360
DB 301 ELLDKYDLSNIVEIASGAPLSKEIGEAVARRENLPGRVGGLTETTSATITPEGDK 360
QY 361 PGASGVVPLFAKAYIDLTQKTKLPNRRGEVCAVPMKGYDNBPATREIIDEQML 420
DB 361 PGASGVVPLFAKAYIDLTQKTKLPNRRGEVCAVPMKGYDNBPATREIIDEQML 420
QY 421 HFGDIGYDEEKHFIVDRKLKIKKGYQVPALESLVLOHPNIFDAGVAVDPDPA 480
DB 421 HFGDIGYDEEKHFIVDRKLKIKKGYQVPALESLVLOHPNIFDAGVAVDPDPA 480
QY 481 ELPGAVVPLKKGKSMTEKVMYVSAOVSAKRLRGVAFVDEVPKGLTGKIDKAIREI 540
DB 481 ELPGAVVPLKKGKSMTEKVMYVSAOVSAKRLRGVAFVDEVPKGLTGKIDKAIREI 540
QY 541 LKRPVAKM 548
DB 541 LKRPVAKM 548

RESULT 13

; US-08-076-042-2
; Sequence 2, Application US/08076042

; Patent No. 5330906

; GENERAL INFORMATION:

; APPLICANT: KAJIYAMA, NAOKI

; APPLICANT: NAKANO, EIICHI

; TITLE OF INVENTION: MUTANT LUCIFERASE OF A FIREFLY, MUTANT

; TITLE OF INVENTION: LUCIFERASE GENES, NOVEL RECOMBINANT DNAs CONTAINING THE

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PENNIE & EDMONDS

; STREET: 1155 AVENUE OF THE AMERICAS

; CITY: NEW YORK

; STATE: N.Y.

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/076,042

; FILING DATE: 15-JUN-1993

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/675,211

; FILING DATE: 26-MAR-1991

; NAME: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: MISTROCK, S. LESLIE

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 7005-026-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-790-9090

; TELEFAX: 212-869-9741

; TELEFAX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 548 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; ORIGINAL SOURCE:

; ORGANISM: Luciola cruciata

; US-08-076-042-2

Query Match 95.1%; Score 2684; DB 1; Length 548;

Best Local Similarity 93.4%; Pred. No. 5.5e-272;
Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

```

OY 1 MEMMENDENIVYGPPEPPIEESGAGOLRKYMDRYAKIGAIAFTNALGVDYTYAYLE 60
Db 1 MEMMENDENIVYGPPEPPIEESGAGOLRKYMDRYAKIGAIAFTNALGVDYTYAYLE 60
OY 61 KSCCLGKALONGLVVDGRIALCSECEEFFIPVLAGLFIGVAVPTNEITYLRELHSL 120
Db 61 KSCCLGKALONGLVVDGRIALCSECEEFFIPVLAGLFIGVAVPTNEITYLRELHSL 120
OY 121 GISKPTIVSSKGLDKVTVQKTAIYIVLDSKVDYRGYQSDNFIKNTPOGFGK 180
Db 121 GISKPTIVSSKGLDKVTVQKTAIYIVLDSKVDYRGYQSDNFIKNTPOGFGK 180
OY 181 SSFKTVEVRKKEQVALINSSSGTGLPKGVOLTHENITVRFSHARDPIYGNVSPGTAIL 240
Db 181 SSFKTVEVRKKEQVALINSSSGTGLPKGVOLTHENITVRFSHARDPIYGNVSPGTAIL 240
OY 241 TVVPHHGFMTTIGYLTGCFRIVMLTKFDEETFLKTLQDYKSSVILVPTLFAILNS 300
Db 241 TVVPHHGFMTTIGYLTGCFRIVMLTKFDEETFLKTLQDYKSSVILVPTLFAILNS 300
OY 301 ELLDKYDLSNLEIASGAPLSKEIGEAARRPNLPVARGYGLTETTSAILITPEGDK 360
Db 301 ELLDKYDLSNLEIASGAPLSKEIGEAARRPNLPVARGYGLTETTSAILITPEGDK 360
OY 361 PGASGRVPLFRAKYIDLDTKKTLGPNRRGEVCVGMIMKGYVNDPEATREIIDEQWL 420
Db 361 PGASGRVPLFRAKYIDLDTKKTLGPNRRGEVCVGMIMKGYVNDPEATREIIDEQWL 420
OY 421 HTGDIGYDEEHFPIVDRKSLIKYKQVPPALESVLLQHPNIFDAGVAGVDDPAG 480
Db 421 HTGDIGYDEEHFPIVDRKSLIKYKQVPPALESVLLQHPNIFDAGVAGVDDPAG 480
OY 481 ELPGAIVVLLKKGSMTEKEVMDYASOVSNARKLRGVRFVDEVPKGLTGKIDGAIREI 540
Db 481 ELPGAIVVLLKKGSMTEKEVMDYASOVSNARKLRGVRFVDEVPKGLTGKIDGAIREI 540
OY 541 LKPPVAKM 548
Db 541 LKPPVAKM 548

```

RESULT 14

US-09-380-061B-14
Sequence 14, Application US/09380061B
Patent No. 6265177

GENERAL INFORMATION:

APPLICANT: SQUIRRELL, DAVID JAMES
WHITE, PETER JOHN
LOWE, CHRISTOPHER ROBIN
MURRAY, JAMES AUGUSTUS HENRY
TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/380,061B
FILING DATE: 25-Aug-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB98/01026

FILING DATE: 7-Apr-1998
APPLICATION NUMBER: GB 9707468.8
FILING DATE: 11-Apr-1997
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B. J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 124-725
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4000
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-380-061B-14

Query Match

Best Local Similarity 95.1%; Score 2684; DB 4; Length 548;
Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

```

OY 1 MEMMENDENIVYGPPEPPIEESGAGOLRKYMDRYAKIGAIAFTNALGVDYTYAYLE 60
Db 1 MEMMENDENIVYGPPEPPIEESGAGOLRKYMDRYAKIGAIAFTNALGVDYTYAYLE 60
OY 61 KSCCLGKALONGLVVDGRIALCSECEEFFIPVLAGLFIGVAVPTNEITYLRELHSL 120
Db 61 KSCCLGKALONGLVVDGRIALCSECEEFFIPVLAGLFIGVAVPTNEITYLRELHSL 120
OY 121 GISKPTIVSSKGLDKVTVQKTAIYIVLDSKVDYRGYQSDNFIKNTPOGFGK 180
Db 121 GISKPTIVSSKGLDKVTVQKTAIYIVLDSKVDYRGYQSDNFIKNTPOGFGK 180
OY 181 SSFKTVEVRKKEQVALINSSSGTGLPKGVOLTHENITVRFSHARDPIYGNVSPGTAIL 240
Db 181 SSFKTVEVRKKEQVALINSSSGTGLPKGVOLTHENITVRFSHARDPIYGNVSPGTAIL 240
OY 241 TVVPHHGFMTTIGYLTGCFRIVMLTKFDEETFLKTLQDYKSSVILVPTLFAILNS 300
Db 241 TVVPHHGFMTTIGYLTGCFRIVMLTKFDEETFLKTLQDYKSSVILVPTLFAILNS 300
OY 301 ELLDKYDLSNLEIASGAPLSKEIGEAARRPNLPVARGYGLTETTSAILITPEGDK 360
Db 301 ELLDKYDLSNLEIASGAPLSKEIGEAARRPNLPVARGYGLTETTSAILITPEGDK 360
OY 361 PGASGRVPLFRAKYIDLDTKKTLGPNRRGEVCVGMIMKGYVNDPEATREIIDEQWL 420
Db 361 PGASGRVPLFRAKYIDLDTKKTLGPNRRGEVCVGMIMKGYVNDPEATREIIDEQWL 420
OY 421 HTGDIGYDEEHFPIVDRKSLIKYKQVPPALESVLLQHPNIFDAGVAGVDDPAG 480
Db 421 HTGDIGYDEEHFPIVDRKSLIKYKQVPPALESVLLQHPNIFDAGVAGVDDPAG 480
OY 481 ELPGAIVVLLKKGSMTEKEVMDYASOVSNARKLRGVRFVDEVPKGLTGKIDGAIREI 540
Db 481 ELPGAIVVLLKKGSMTEKEVMDYASOVSNARKLRGVRFVDEVPKGLTGKIDGAIREI 540
OY 541 LKPPVAKM 548
Db 541 LKPPVAKM 548

```

RESULT 15

US-09-111-752-7
Sequence 7, Application US/09111752
Patent No. 6074859

GENERAL INFORMATION:
APPLICANT: HIROKAWA, KOZO
APPLICANT: KAJIYAMA, NAOKI
APPLICANT: MURAKAMI, SEIJI

TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
 TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE BIOLUMINESCENT PROTEIN
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
 CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/111,752
 FILING DATE: 08-JUL-1998
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 552 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Luciola cruciata and Phantius pyralis
 US-09-111-752-7

Query Match 91.7%; Score 2590; DB 3; Length 552;
 Best Local Similarity 91.0%; Pred. No. 3.9e-262;
 Matches 494; Conservative 31; Mismatches 18; Indels 0; Gaps 0;

QY 1 MENNENDENIVYGPPEPPYPIEGSSAGAOQLRRKYMRYAKLGAIAFTNALTGVDYTYAEYLE 60
 DB 1 MENNENDENIVYGPPEPPYPIEGSSAGAOQLRRKYMRYAKLGAIAFTNALTGVDYTYAEYLE 60
 QY 61 KSCCLGALKNYGLVYDGRITALCSNCEEFFIPYLAGLFTIGVGYAPTNETYTLRELYHSL 120
 DB 61 KSCCLGALKNYGLVYDGRITALCSNCEEFFIPYLAGLFTIGVGYAPTNETYTLRELYHSL 120
 QY 121 GISKPTIVFSSKKGIDVITVOKTVTAIKTIVILDSVDYRGYOSMDNFKKNTPOGPKG 180
 DB 121 GISKPTIVFSSKKGIDVITVOKTVTAIKTIVILDSVDYRGYOSMDNFKKNTPOGPKG 180
 QY 181 SSFRTVEVNRKEQVALIMSSGSTGLPKGVQLTHENIVTRFSHARDPIYGNQVSPGTAIL 240
 DB 181 SSFRTVEVNRKEQVALIMSSGSTGLPKGVQLTHENIVTRFSHARDPIYGNQVSPGTAIL 240
 QY 241 TVVPFHGFGMTTGLVTCGFRIVMLTKPDETEFLKTLQDYKCSYILVPTLFAILNRS 300
 DB 241 TVVPFHGFGMTTGLVTCGFRIVMLTKPDETEFLKTLQDYKCSYILVPTLFAILNRS 300
 QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGYGLTETTSAILITPEGDDK 360
 DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAVARFNLPGVROGYGLTETTSAILITPEGDDK 360
 QY 361 PGASGVVPLFFAKVIDLDTKTKLGNRRGEVCVKPMLKMGVYDNPDEATREIIDEQWL 420
 DB 361 PGASGVVPLFFAKVIDLDTKTKLGNRRGEVCVKPMLKMGVYDNPDEATREIIDEQWL 420
 QY 421 HTGDIGYDEEKHFETIDRLKSLIKKGYQVPAELESVLLQHPNIPDAGVAGLPDDPDIAG 480
 DB 421 HTGDIGYDEEKHFETIDRLKSLIKKGYQVPAELESVLLQHPNIPDAGVAGLPDDPDIAG 480

Db 421 HTGDIGYDEEKHFETIDRLKSLIKKGYQVPAELESVLLQHPNIPDAGVAGLPDDPDIAG 480
 QY 481 ELLPGAVVYLLKKGKSMTEKEVMDYVASOYNAKRLRGVRFYDVEVKGILTKGIDKAIRI 540
 Db 481 ELLPAVVVLEHGKTMTEKEIVDYVASOYTTAKKLGGVVFDEVPKGLTKIDARKIRI 540
 QY 541 LKK 543
 Db 541 LKK 543

Search completed: November 29, 2001, 01:53:48
 Job time: 3208 sec


```

? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/07/903,047
? FILING DATE: 19920623
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Mistrock, S. Leslie
? REGISTRATION NUMBER: 18,872
? REFERENCE/DOCKET NUMBER: 7005-048
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 212 790-9090
? TELEFAX: 212 869-8864/9741
? TELEX: 66141 PENNIE
? INFORMATION FOR SEQ ID NO: 7:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1644 base pairs
? TYPE: NUCLEIC ACID
? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: cDNA to mRNA
? US-07-903-047-7

```

```

alignment_scores:
  Quality: 2814.00      Length: 548
  Ratio: 5.144         Gaps: 0
  Percent Similarity: 99.818      Percent Identity: 99.635

```

```

alignment_block:
US-09-581-241-6 x US-07-903-047-7

```

```

Align seg 1/1 to: US-07-903-047-7 from: 1 to: 1644

```

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProph 17
1 ATGGAACAAATGAGACGATGAAATATGCTGATGCTCTGGAACCATTT 50
17 eTyrProIleGluGlySerAlaGlyAlaGluLeuArgLysTyrMetC 34
51 TTACCTATTTGAAAGAGATGCTGAGACCAATTCGCAAGTATATGCG 100
34 sPARGTrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 ATCGATATCAAACTTGAGCAATTCCTTTACTAACGCACTTACCGGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGlyLysSerCysLeuGlyG 67
151 GTCGATTATACGTACGCCAATATTAGAAAATCATGCTGTCTAGAGA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
201 GCGTTTAAAGAAATATGCTTGTGATGAGAAATTCGCTATATGCA 250
84 eArgLysAsnCysGluPhePheIleProValLeuAlaGlyLeuPheI 100
251 GTGAAAACAGTGAAGAAATCTTATTCCTGATTAAGCCGTTATTTA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeu 117
301 GGTTGCGGTGGCTCCAACTAATGATTTACACTCTACGTAAGTAATG 350
117 IHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
351 TCACATTTTAGCATCTCTAAGCAACATTTGATTTAGTTCTCAAAAAG 400
134 ILeuAspLysValIleThrValGlyLysThrValThrAlaIleLysTr 150
401 GATTAGTAAGTTAATACGTACAAAAACGTAACGTAATTAATAAAC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
451 ATTGTTATTTGACAGCAAGATGATATATAGGTTATCATCATTCATGA 500

```

```

167 pAsnPheIleLysAsnThrProGlnGlyPheLysGlySerPheI 184
501 CAACCTTATTAATAAAACACCTCCACAGGTTTCAAGCATCAAGTTTA 550
184 ySerValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAACCTAGAGTTAACCCGCAAGAACAGTTGCTCTTATTAATGAACCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
601 TCGGGTTCAACCGTTTCCCAAAACCTGTGCAACTTACATCAAAAATC 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
234 eProGlyThrAlaIleLeuThrValValProPheHisIleGlyPheGly 250
701 CACAGGACACGCTATTTTAAGTATGATACATTCATCATGCTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMet 267
751 ATGTTTACTACTTATGAGCTATCTAAGTGTGCTTGTGATTTGCTATG 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
801 AACCAATTTGACGAAAGACACTTTTAAAAACACGCAAGATACAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAACGCTTATTTCTGTACCGCACTTGTGTCATTTCTTAATAAAG 900
301 GluLeuAspLysTyrAspLeuSerAsnLeuValGluIleLasergl 317
901 GAATTAAGTATGATATGATTTATCAAAATTTAGTCAAAATTCATCTG 950
317 yGlyAlaProLeuSerLysGluIleGlyAlaValAlaArgArgPhe 334
951 CGGAGCACTTATCTAAGAAATGCTGAAGCTGTGCTAGACGTTTGA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrTrpSer 350
1001 ATTTACCGGGTGTGTCAGAGCTATGCTTTAACAGAAACACCTCTGCA 1050
351 IleIleIleThrProGlnGlyAspAspLysProGlyLaserglyLys 367
1051 ATTATTTATCACCGGAAGCGGATGATTAACCAAGCTCTCTGCAAAAG 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThr 384
1101 TGTCCATTTATTAAGCAAAAGTTATCGATCTTGATCAAAAAAATT 1150
384 euGlyProAsnArgArgGlyValValCysValLysGlyProMetLeuMet 400
1151 TGGGCCGCAACGACGTGAGAGATTTGTAAAGGCTCTATCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgLysIleLeuAspGlu 417
1201 AAAGGTTATGTAGATTAATCCAGAAAGCAACAGCAAAATCATAGTAGA 1250
417 uGlyThrPheLysThrGlyAspIleGlyTyrTyrAspGluLysHisP 434
1251 AGGTTGTTGACACAGAGATATTGGTATTAACATGAAACAAAACATT 1300
434 hepHeIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTATCTGTCGATCTTGAAGTCTTTAATCAAAATACAAAGGATACAA 1350
451 ValProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePhe 467
1351 GTACACCTGCTGAATTAGAATCTGCTTTTGCMAACATCCAAATATTTT 1400

```

```

467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
1401 TGAATGCCGGCTGCTGGCTCCAGATTCCTATATGCTGGAGCTTCGG 1450
484 LyAlaValAlaValLeuLysGlyGlySerMetThrGluLysGluVal 500
1451 GAGCTGTGTGTACTTGAAGAAAGAAATCTATGACTGAAAAAGAGTA 1500
501 MetAspTyrAlaAlaSerGluValIleSerAlaLysArgLeuArgGly 517
1501 ATGATTAAGTCTGCTAGTCAAGTTCAATGCAAAACGTTGGCTGGTGG 1550
517 yValaArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
1551 TGTCCGCTTTTGTGACGAAGTACCTAAAGSTCTCAGCTGTAATAATGACG 1600
534 LyysAlaIleArgGluIleLeuLysGlySerProValAlaLysMet 548
1601 GTAAACCAATTAGAGAAATCTGAGAAACCAAGTGTCTAAGATG 1644

```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:us-09-380-061B-15

seq_documentation_block:

Sequence 15, Application US/09380061B
Patent No. 6265177

GENERAL INFORMATION:

APPLICANT: SOUIRRELL, DAVID JAMES

WHITE, PETER JOHN

LOWE, CHRISTOPHER ROBIN

MURRAY, JAMES AUGUSTUS HENRY

TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESS: NIXON & VANDERHUYE P.C.

STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/380,061B

FILING DATE: 25-Aug-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB98/01026

FILING DATE: 7-APR-1998

APPLICATION NUMBER: GB 9707468.8

FILING DATE: 11-APR-1997

ATTORNEY/AGENT INFORMATION:

NAME: SADOFF, B. J.

REGISTRATION NUMBER: 36,663

REFERENCE/DOCKET NUMBER: 124-725

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)816-4100

TELEFAX: (703)816-4100

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 1644 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1644

SEQUENCE DESCRIPTION: SEQ ID NO: 15:

US-09-380-061B-15

alignment_scores:

Quality: 2814.00 Length: 548

Ratio: 5.144 Gaps: 0

Percent Similarity: 99.818 Percent Identity: 99.635

alignment_block:

US-09-581-241-6 x US-09-380-061B-15 ..

Align seq 1/1 to: US-09-380-061B-15 from: 1 to: 1644

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProG 17
1 ATGGAAGAAACATGGAAGACGATGAAATATTTGATGCTCTGAAACATT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMet 34
51 TTACCCCTATTGAAGAGGATCTGCTGGAGCACAATTGCCAGATTATGG 100
34 sParGTYrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 ATCGATATGCAAAACCTGGAGCAATTCCTTACTTAACGCACTTACCGGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
151 GTGATTTATATGCTACGCCGCAATCTTAGAAAAATCATGCTGCTTAGACA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgGlyAlaLeuCys 84
201 GGCTTTAAAGAAATTATGTTGGTGTGTGATGAGAAATGCTTATGCA 250
84 eArgLysnGlyGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAAGCTGTGAAGATCTTATTCCTGATTACCGGTTATTATTAA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeu 117
301 GGTGCTGCTGTGCTCCAACTATGAGATTTCACCTACGTAATTTGGT 350
117 HisSerLeuGlyIleSerLysProThrIleValPheSerLysLysG 134
351 TCACATTTAGGATCTTACGCAACAATGTAATTTACTTTCAAAAAAG 400
134 LyLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
401 GATTAGATTAAGTTAATCTGACAAAAAACGTAATCTCTTTAAAC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
451 ATTGTTATATTGACAGCAAGGTGATTATGAGGTTATCAATCCATGCA 500
167 PAsnPheIleLysLysAsnThrProGlnIlePheLysGlySerSerPhe 184
501 CAACCTTTATTAATAAAAAACACTCCCAAGTTTCAAGGATCAAGTTTA 550
184 ySThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAACGTGAGAAAGTTAAACGCAAGAAAGAGTTGCTTATATGAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
601 TCGGGTTCACCGGTTTGCCAAAAGGTGTCCAACTTACTCAAGAAATGC 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
651 AGTACATGATTTTCTCAGCGCTAGAGATCCCAATTATGAAACCAAGTTT 700
234 eTProGlyThrAlaIleLeuThrValValProPheHisIleGlyPheGly 250
701 CACGAGCAGCGCTATTTTAAGTATGATGATCCATCATCATGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheAlaGlyIleValMet 267

```

```

751 ATGTTACTACTTACGCTATCTACTTGTGTTTGTATGTCATGTT 800
267 uThrLysPheAspGluuThrPheLeuLysThrLeuGlnAspTyrIleSc 284
801 AACGAAATTCACGAGACACTTTTAAAAACACTGCAGATTACCAAT 850
284 ysserSerValIleLeuValProthrLeuPheAlaIleuAsnArgSer 300
851 GTTCAAGCGTATCTTGTACCGACTTGTGTCATTGTCATTCTAATAGAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
901 GAATTACTCGATTAATATGATTATTCAAATTACTGTAATTCATCTCG 950
317 yGlyAlaProLeuSerLysGluIleGlyAlaValAlaArgArgPhe 334
951 CGGAGCACCCTTATCTAAGAAATGTTGTAAGCTGCTAGACGTTTAA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAla 350
1001 ATTTACCGGGTGTGTCAGAGCTATGTTTACAGAAACACTCTTGCA 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
1051 ATTATTATCACACCGAGGCGGATGATTAACCGAGTCTTGCCAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
1101 TGTGCAATTATTTAAAGCAAAAGTTATCGATCTGATCACTAAAAAAGCTT 1150
384 euGlyProAsnArgArgLysGlyLysValLysGlyProMetLeuMet 400
1151 TGGGCCCCACAGAGCTGAGAGAGTTTGTAAAGGCGCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleLeuAspGlu 417
1201 AAAGGTATGTAGATAATCCAGAACCAAGAGAAATCATAGATGAAGA 1250
417 uGlyTrpLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysH 434
1251 AGCTGGTGGCACACAGAGATTTGGTATTCATCATAGACAAAAACATT 1300
434 hepHeIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTATCTGATGATGCTTTGAAGCTTTAATCAATACAAAGATATCAA 1350
451 ValProProAlaGluLeuGlnSerValLeuLeuGlnHisProAsnIleP 467
1351 GTACCACTGCTGAATTAATCTGCTTTGCAACATCCAAATATTTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuPro 484
1401 TGATGCCGGCGTGTGCTGGCGTTCCKAGATCTTACGTGTGAGCTTCCG 1450
484 lValAlaValAlaValLeuLysLysGlyLysSerMetThrGlnLysGluVal 500
1451 GAGCTTGTGTTGACTGTAAGAAAGAAATCTATGACTCAAAAGAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
1501 ATGGATTAGCTGTAGTCACAGTTTCAAATGCAAAACGTTTCCGTGGTG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
1551 TGTCCTGTTTGTGACGAGAGTACTAAAGTCTCACTGTGTAATATGAG 1600
534 lYlYsAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
1601 GTAAGCAATTAGAGAAATCTGAGAAACCAAGTTCTTAAGATG 1644
seq_name: /cgn2.6/ptodata/2/1na/5A_COMB.seq:us-08-460-934-8
seq_documentation_block:

```

```

: Sequence 8, Application US/08460934
: Patent No. 5814465
:
: GENERAL INFORMATION:
: APPLICANT: TATSUMI, HIROKI
: APPLICANT: FUKUDA, SATOSHI
: APPLICANT: KIKUCHI, MAMORU
: APPLICANT: KOBAYASHI, YASUJI
: TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
: TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
: TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
: ANALYSTS METHOD
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: P. C.
: STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
: CITY: ARLINGTON
: STATE: VA
: COUNTRY: USA
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/460,934
: FILING DATE: 05-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 193798/1994
: FILING DATE: 27-JUL-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 54625/1995
: FILING DATE: 14-MAR-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 98857/1995
: FILING DATE: 24-APR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: OBLON, NORMAN F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 7126-001-0
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-413-2220
: TELEFAX: 703-413-3000
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1908 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: misc. feature
: LOCATION: 1..1908
: OTHER INFORMATION: /note="The nucleotide sequence of
: the biotinylated firefly luciferase gene contained in
: recombinant plasmid pHLf248 DNA"
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1908
: US-08-460-934-8
:
: alignment_scores:
: Quality: 2812.00 Length: 547
: Ratio: 5.141 Gaps: 0
: Percent Similarity: 100.000 Percent Identity: 99.634
:
: alignment_block:
: US-09-581-241-6 x US-08-460-934-8
:
: Align seg 1/1 to: US-08-460-934-8 from: 1 to: 1908

```

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProph 17
1 ATCGAAACACGTGAGAACGATGAAATATTGTGTATGCTTACCAACCAT 50
17 eYrProIleGlyGlySerIleGlyValGluLeuArgLysTyrMet 34
51 TTACCTATTGAAGAGGATCTGTGTAGACACAATTGCCAAGTATATGG 100
34 sParGThrIleLysLeuGlyValIleAlaPheThrAsnAlaLeuThrGly 50
101 ATCGATTATGCAAACTTGGAGCATTTGCTTACTAAGCATCTTACCGGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGlyLysSerCysCysLeuGly 67
151 GTGAGTATTACGACCGCAATCTAGAAAATCATGGCTGTAGAGAGA 200
67 ValLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
201 GGCTTTAAAGAAATTATGCTTGTGTGTATGGAAGAAATTCCTTATGCA 250
84 eRGluAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAACACGTGAGAGTCTTATTCTGTATTTACCGGTTTATTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeu 117
301 GGTTGCGGTGGCTCCAACTATGAGATTTACACTCTACGTCAATTTGGT 350
117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
351 TCACAGTTTAGCATCTTACGCCAACAAATGATTTAGTTCTTAAAGAA 400
134 ILeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
401 GATTAGATTAAAGTATTAAGTACCAAAAACGTAACCTCTATTAAAGC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
451 ATTGTTATATTTGGACAGCAAAAGTATTAAGTATTAAGTATTAAGT 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
501 CAACCTTTATTAATAAAACACCTCCAAAGGTTTCAAAAGATCAAGTTTA 550
184 YSThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAACGTAGAAAGTAAACCGCAAAAGAAAGTTGCTTATATGAACCTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
601 TCGGTTTACCGGTTTCCCAAAAGGTTGCAACTTACTCATGAAATTTT 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
651 GGTACACGGGTTTTCTCACGCTAGAGATCCAAATTTATGAAACCAAGTTT 700
234 eRPProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
701 CACCAAGGACGGCTATTTTAACTAGTACCATTCATCATGCTTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
751 ATGTTTACTTATAGCTATCTAATCTGTGTTTTCGTATTTGATATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLys 284
801 AACGAAATTTGACGAAGAGACTTTTAAACACTGCACAGATTCACAAT 850
284 YsSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCNAGCGTTATCTGTACGACTTTGTTGCAATTCCTAATATAGAAGT 900

```

```

301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
901 GAATTTACTGATTAATATGATTTATCAAAATTTAGTTGAAATTCATCTGG 950
317 YGlyAlaProLeuSerLysGluIleGlyValAlaValAlaArgArgPhe 334
951 CGGAGACACCTTATCTAAAGAAATTTGTTGAAGCTGTTGCTAGACGTTT 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
1001 ATTTACCGGGTGTTCGTCACAGGCTATGGTTTAAACAAGAACCTCTGCA 1050
351 IleIleIleThrProGluGluAspAspLysProGlyValSerLysVal 367
1051 ATTATTTATCACCGGAGGCGATGATTAACGAGGCTTCTGCAAAAGT 1100
367 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
1101 TGTGCATTATTAAAGCAAAAGTATTCGATCTTGATTAATAAAACCTT 1150
384 eUGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
1151 TGGGCCCGAAGACGCGTGAAGATTGTTGTAAGGGCTCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
1201 AAAGTTATGATGATATATCCAGACAGACAGAAATATCATGATATGAGA 1250
417 uGlyThrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHis 434
1251 AGTTGGTTGCACACAGGAGATTTGGTATTAAGTAAAGAAACAACTT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTTATCTGATCTGTTGAAGCTTTTATATCAAAATCAAGGATATCA 1350
451 ValProProAlaGluLeuGlnSerValLeuLeuGlnHisProAsnIlePh 467
1351 GTACACCTGCTGAATTAAGATCTGTTTGGCAACATCCAAATATTTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuPro 484
1401 TGATGCGGCGTTGCTGGCGTTCCAGATCTTATAGCTGGTGAAGCTCCGG 1450
484 IValAlaValAlaLeuLysLysGlyLysSerMetThrGluLysGluVal 500
1451 GAGCTGTTGTTGCTGTAAGAAAGAAATCTATGACTCAAAAAGAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
1501 ATGATTTACGTTGCTAGTCAAGTTCAAAATGCAAAACGTTTCTGCTGTCG 1550
517 YValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
1551 TGTCCGTTTGTGGAGCAAGTACCTTAAGGTCTCACTGGTGAATAATGACG 1600
534 IYValAlaIleArgGluIleLeuLysLysProValAlaLys 547
1601 GTAAAGCAATTAGAAATACTGAGAAACAGGTTGCTAG 1641
seq_name: /cgn2_6/ptcdat2/ina/5b_comb.seq:us-08-782-118-8
seq_documentation_block:
: Sequence 8, Application us/08782118
: Patent No. 5843746
: GENERAL INFORMATION:
: APPLICANT: TATSUMI, HIROKI
: APPLICANT: FUKUDA, SATOSHI
: APPLICANT: KIKUCHI, MAMORU
: APPLICANT: KOYAMA, YASUJI
: TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
: TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A GENE
: TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT

```



```

? TITLE OF INVENTION: ANALYSIS METHOD
? NUMBER OF SEQUENCES: 14
? CORRESPONDENCE ADDRESS:
? ADDRESS: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
? ADDRESSEE: P.C.
? STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
? CITY: ARLINGTON
? STATE: VA
? COUNTRY: USA
? ZIP: 22202
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/782,118
? FILING DATE: 13-JAN-1997
? CLASSIFICATION: 435
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 08/460,934
? FILING DATE: 05-JUN-1995
? APPLICATION NUMBER: JP 193798/1994
? FILING DATE: 27-JUL-1994
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: JP 54625/1995
? FILING DATE: 14-MAR-1995
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: JP 98857/1995
? FILING DATE: 24-APR-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: OBLON, NORMAN F.
? REGISTRATION NUMBER: 24,618
? REFERENCE/DOCKET NUMBER: 7126-001-0
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 703-413-3000
? TELEFAX: 703-413-2220
? INFORMATION FOR SEQ ID NO: 8:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1908 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: unknown
? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: misc.feature
? LOCATION: 1..1908
? OTHER INFORMATION: /note= "The nucleotide sequence of
? OTHER INFORMATION: the biotinylated firefly luciferase gene contained in
? OTHER INFORMATION: recombinant plasmid pHLf248 DNA"
? NAME/KEY: CDS
? LOCATION: 1..1908
? US-08-782-118-8

```

```

alignment_scores:
  Quality: 2812.00      Length: 547
  Ratio: 5.141          Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 99.634

```

```

alignment_block:
  US-09-581-241-6 x US-08-782-118-8  ..
  Align seg 1/1 to: US-08-782-118-8 from: 1 to: 1908

```

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTYRGLYProGluProph 17
1 ATGGAAGACATGAGACGATGAAATATGTGTATGCTGACCAACATT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnIleuArgLysTyrMetA 34
51 TTACCTATTGAAGAGACTGCTGAGACACAAATTGCCAAGTATATAG 100

```

```

34 spargTyfAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 ATCGATATCCAAACTGGAGCAATTCCTTTACTACCACTTACCGCT 150
51 ValAspTyrThrTyraAlaGluTyrLeuGluLysSerCysGlyGlyGly 67
151 GTGGATTATACGACCGCAATCTAGAAAATCATGCTGCTGAGAGA 200
67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuCys 84
201 GGCTTTAAAGAAATTATGCTTGTGTTGATGGAAGAAATGCGTATGCA 250
84 ecGluAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAACTGTGAGAGTCTTATTCCTGATATACCGCTTATTTANA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVa 117
301 GTGTCCGCTGTGCTCCAACTAATGAGATTACACTGCTGGAATTGCT 350
117 IHisSerLeuGlyIleSerLysProThrIleValPheSerLysLysG 134
351 TCACAGTTTAGCATCTCTAAGCCAAACAATGTATTAGTTCTAAAGAG 400
134 ILeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
401 GATTAGTAAAGTTATTAAGTCAAAAAAAGCGTACTGCTATTAAAGC 450
151 ILeuValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
451 ATTGTATATTGGACAGCAAGTGAGATTATAGGTTATCAACCATGGA 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
501 CAAGTTATTAAATAAAACACTCCACAGAGTTTCAAGATCAAGTTTA 550
184 ySThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAAGTATAGAAAGTTAACCCGCAAGAACAGTCTGCTATTAAAGAACTG 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
601 TCGGTTCAACCGGTTGCCAAAGGTGTCACACTCACTATGAAAATTT 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
651 GGTACAGCGTTTCTCAGCGTACAGATCCATTTATGGAACCAAGTTT 700
234 erProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
701 CACACAGCAGCGTATTTAACTAGTACCATTCATCATGCTTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMet 267
751 ATGTTACTACTTAGGCTATCTACTTGTGCTTTGCTTATGTCATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
801 AACGAAATTTGACGAGAGACTTTTAAACACCTCAAGATATACAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAAGCGTTATCTGTACCGACTTGTGTCATTTCTTAATAGAACT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
901 GAATTACTCGAATAATATGATTTCATCAATTTAGTGAATTTGCATCTGG 950
317 yGlyAlaProLeuSerLysGluIleGlyAlaValAlaArgArgPheA 334
951 CGAGACACCTTATCTAAAGAAATGTTGGAAGCTGTTCTGAGAGTTT 1000

```

```

334 snleuproglyvalarglnglytyrlyleuthrthrserala 350
1001 ATTACCGGTGTCGTCAGGCTATGCTTACAGAACACCTCTGCA 1050
351 llellelethrprogluglyaspaplysproglyalaserglylsva 367
1051 ATTATATACACCGGAGGCGATGATTAACCGAGTCTTGCGCAAGT 1100
367 lvalprolephelysalaalysvalilleaspleuasptlhrlysthl 384
1101 TGTGCATTATTAAAGCAAACTTTCATCTGTGATCTTAAAAAATCT 1150
384 euglyproasnargarglygluvalcysvallysglyprometleumet 400
1151 TGGGCCGGAACAGCGTGGAGAGTTTGTAAAGGCTCTATGCTTANG 1200
401 lysglytyrvalaspasproglualatrrargluilleaspglucl 417
1201 AAAGTTATGTAGATATCCAGAGCAACAGAAATCATGATGAGACA 1250
417 ucllytrpleuhsithrglyaspilleglytyrtyraspgluglyshsp 434
1251 AGGTGTGTGCACACAGAGATATTGGTATTACGATGAGAAAAACATT 1300
434 hephelelevalasparglyleuylserleullestlytyrlygln 450
1301 TCTTATCTGTGATCGTTGAAGTCTTTAATCAATACAAAGATATCA 1350
451 valproprolagluleugluservalleuenglnhisproasnilleph 467
1351 GTACACCGCTGCAATTAACCTCTTTTGCACATCCAAATATTTT 1400
467 easpalaglyvalalaglyvalproaspptrollealaglyleuprog 484
1401 TGATCCGCGCTTGCTGGCTCCAGATCTATAGCTGCTGCTCCGG 1450
484 lvalalavalalleuylslysglyssermetthrghlulsgluval 500
1451 GAGCTGTTGTCTACTGAAAAAGAAATCTATAGCTGAAAAAGAACTA 1500
501 metasptryvalalaserghlvalserasnaalysargleuarglygl 517
1501 ATGGATTCGCTGTGCTCAAGTTTCAATGCAAAACGTTTGGCGTGG 1550
517 yvalargphevalaspgluvalprolysglyleuthrghlystllespg 534
1551 TGTCCGTTTGTGGAGAGTACCTAAAGGTCTCAGTGTAAATGTGACG 1600
534 llylsalalaleargluilleuylslysprovalalalys 547
1601 GTAAAGCAATTAGAGAAATACTGAAGAAACAGTTGCTTAAG 1641

seq_name: /cgn2_6/protdata/2/ina/5a_COMB.seq:US-08-460-934-5
seq_documentation_block:
; Sequence 5, Application US/08460934
; Patent No. 5814465
; GENERAL INFORMATION:
; APPLICANT: TATSUMI, HIROKI
; APPLICANT: FUKUDA, SATOSHI
; APPLICANT: KIKUCHI, MAMORU
; APPLICANT: Koyama, Yasuji
; TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE
; TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A
; TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VA

```

```

COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,934
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 193798/1994
FILING DATE: 27-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 54625/1995
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 98857/1995
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: OBION, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-001-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: unknown
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1704
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..1704
OTHER INFORMATION:
OTHER INFORMATION: biotinylated firefly luciferase gene contained in recombin
US-08-460-934-5
OTHER INFORMATION: plasmid pHLf203 DNA

alignment_scores:
Quality: 2799.00 Length: 546
Ratio: 5.126 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.267

alignment_block:
US-09-581-241-6 x US-08-460-934-5 ..
Align seg 1/1 to: US-08-460-934-5 from: 1 to: 1704
3 Asnmetgluasnaspgluasnillevaltyrghlyprogluprophe Tyr 19
67 AGCTCGAGACGATGAAATATTTGTGATGTCTCTGAAACCACTTTTACC 116
19 clllegluglugsyeralaglyalaglnleuarglystlyrmetaspargt 36
117 TATGTAAGAGGATCTGCTGGAGACAAATTCGCAAGTATATGATCAT 166
36 yvalalysleuglyalalealaphethrasnalaaleuthrghlyalasp 52
167 ATGCAGAACTTGAGCAATGCTTTTACTTAACGCACTTACCGGTGCGAT 216
53 tyrthrtyralaglytyrleuglulyssercyscysleuglyglualale 69
217 TATAGTACGCCGAGTACTTAGAAAAATCATGCTGTCTAGAGGAGCTTT 266
69 ulysasntryglyleuvalvalaspglyllyrghlyleuylslyserglna 86

```



```

1267 TATGTAGATATCCAGACAGACAGAAATCATATGATGACAGCTTG 1316
419 PLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePheI 436
1317 GTTGACACACAGAGATATGGTATACGATGAGAAACATTTCTTTA 1366
436 LeValAspArgLeuLysSerLeuIleLysTyrIleGlyValPro 452
1367 TCGTGATCGTTGAAGTCTTTATCAAAATCAAGATATCAAGTACCA 1416
453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAl 469
1417 CCGCTGATTAATAGATCTGTTCTTTTGCACATCCAAATATTTTGTATGC 1466
469 aglyValAlaGlyValProAspProIleAlaGlyLeuProGlyAlaVal 486
1467 CGGGGTGCGGGGCTCCAGATCCTATAGCTGTGAGACTCCGGAGCTG 1516
486 aIValValLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
1517 TTGTTGACTTGAAAAAGAAAAATCTATGACTGAAAAAGAAATGCGAT 1566
503 TyrValAlaSerGluValSerAsnAlaLysArgLeuArgGlyValAla 519
1567 TACGTTCTAGTCAAGTTTCAATGCAAAACGTTTGGCTGTGTGCTCCG 1616
519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA 536
1617 TTTGTGAGACAGTACCTTAAGCTCTCAGCTGTAATTAATGACGGTAAAG 1666
536 IaIleArgGluIleLeuLysLysProValAlaLysMet 548
1667 CAATTAGAAATCTGAGAAACCAAGTCTGTAAGATG 1704

seq_name: /cgn2_6/plodata/2/lna/5A_COMB.seq:US-07-675-211-1
seq_documentation_block:
; Sequence 1, Application US/07675211
; Patent No. 5219737
; GENERAL INFORMATION:
; APPLICANT: KAJIYAMA, NAOKI
; APPLICANT: NAKANO, EIICHI
; TITLE OF INVENTION: MUTANT LUCIFERASE OF A FIREFLY, MUTANT
; TITLE OF INVENTION: LUCIFERASE GENES, NOVEL, RECOMBINANT DNAs CONTAINING THE
; TITLE OF INVENTION: GENES AND A METHOD OF PRODUCING MUTANT LUCIFERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: N.Y.
; COUNTRY: U.S.A
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/675,211
; FILING DATE: 19910326
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISHROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7005-026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1644 base pairs

```

```

; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Luciola cruciata
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1644
; US-07-675-211-1

alignment_scores:
    Quality: 2684.00      Length: 548
    Ratio: 4.943          Gaps: 0
    Percent Similarity: 99.088      Percent Identity: 93.431

alignment_block:
US-09-581-241-6 x US-07-675-211-1 ..
Align seg 1/1 to: US-07-675-211-1 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
1 ANGGAAGACATGGAAGAAACGATGAAATATTTAGTGGACCTAAACCGTT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetA 34
51 TTACCTATCGAAGAGGATCTGCTGGAACACAAATTAACGCAATACATGG 100
34 sParGTrAlaLysLeuGluAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 AGCGATATCGAAACACTTGGCGCAATGCTTTACAAATCGACTGTGCT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGly 67
151 GTGATATATCTTACCGCGAATCTTGAGAAATCATGTTGCTAGAGAA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgGluAlaLeuCys 84
201 AGCTTGCAAAATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 250
84 eArgLysCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAGCTGGAAGATTTTATTCGTGATATACCGGACTGTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgLeuVal 117
301 GGGTAGCTGTGACCCCACTAATGAGATTTCACCTTACGTAAGTGGT 350
117 HisSerLeuGlyIleSerLysProThrIleValIleSerSerLysGly 134
351 TCACAGTTTAGGATCTCTAAACCAACAAATTTAGTTCAAAAAG 400
134 LysAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
401 GCTTAGATAAAGTATTAACAGTACAGAAACAGTACTATTAATAACC 450
151 IleAlaIleLeuAspSerLysValAspTyrArgIleTyrGlnSerMet 167
451 ATTGTATACATGACCAAAAGTGAATATGAGAGATATCAAGTCTGCA 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
501 CACCTTTATTAAGAAACACTCCACAGGTTTCAAGCATCCAGTTTCA 550
184 yStrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
601 TGGGTTCTACCGGTTTGCCAAAAGGCGTACACTTACTACAGAAATAC 650

```

```

217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
|||||
651 AGCAGCTAGATTTCATGCTGAGATCGATTTATGTAACCAACTTT 700
234 ePProGlyThrAlaIleLeuThrValValProPheHisGlyPheGly 250
|||||
701 CACCGACGACCGCTGTTTAACTGCTCCATTCATCATGCTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMet 267
|||||
751 ATGTCACACTCTAGGCTATTTAATTTGGGTTTCGTTGTTGTAAGTT 800
267 uThrIysPheAspGluGluThrPheLeuIysThrLeuGlnAspTyrLysC 284
|||||
801 AACCAAAATTCGATGAGAAGAAACATTTTAAACTCTACAAAGATTATAAT 850
284 ySSerSerAlIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
|||||
851 GTACAAAGTGTATCTTGTGACGACCTGTTTGCATTCCTCAACAAAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
|||||
901 GAATTAACCTCAATAAATACGATTGTCATATTTAGTTGACATTCGATCTG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
|||||
951 CGGACGACCTTATCAAAAGAGTGTGACGCTGTTGCTACGACGCTTAA 1000
334 snLeuProGlyValAlaArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
|||||
1001 ATCTCCCGGCTGCTGCTCAGGTTATGCTTAAACGAAACCAACATCTGCC 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
|||||
1051 ATTATTATTACACGAGAGACGATTAACCAAGAGCTTGGAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrIysLysThrL 384
|||||
1101 CGTGCCGTTGTTTAAAGCAAAAGTTATGATCTTATACCAAAAATCTT 1150
384 euGlyProAsnArgArgLysGluValLysValLysGlyProMetLeuMet 400
|||||
1151 TAGGCTCAACGACGAGTGAAGATTGTTAAAGACCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluG 417
|||||
1201 AAAGTATATGTAATAATCAAGAACCAAAAGAACTTATGACGAGA 1250
417 uGlyTyrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
|||||
1251 AGGTTGGCTGCACACGAGATATTGGATTATGATGAGAAAACATTT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
|||||
1301 TCTTATATGTCGATCGTTGAAGCTTTAATCAATACAAAGATACCAA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
|||||
1351 GTACCACTGCCGAATTAAGAAATCGTTCTTTTGCACACATCATCTATCTT 1400
467 eaSPAlaGlyValAlaGlyValProAspProIleAlaGlyLysLeuProG 484
|||||
1401 TGATGCTGGTGTGCGCGCTTCTGATCCTGTAGCTGCGCAGCTTCCAG 1450
484 lYAlaValAlaValLeuLysGlyLysSerMetThrGluLysGluVal 500
|||||
1451 GAGCGCTGTTGTAATGAGAAAGCGGAAATAATATGACCGAAAGAAAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG 517
|||||
1501 ATGGAATTATGTTGCAAGTCAAGTTCAAAATGCAAAACGTTTACGTTGG 1550

```

```

517 yValArgPheValAspGluValProIysGlyLeuThrGlyLysIleAspG 534
|||||
1551 TGTTCGTTGTTGATGAGAACTTAAGCTTAAGCTTACTGGAATAATGACG 1600
534 lYsAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
|||||
1601 GCAGAGCAATTAGAGAAATCTTAAGAAACCACTTGTCTAAGATG 1644
seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-07-903-047-1
seq_documentation_block:
; Sequence 1, Application US/07903047
; Patent No. 5229285
; GENERAL INFORMATION:
; APPLICANT: Kajiyama, Naoki
; APPLICANT: Nakano, Eiichi
; TITLE OF INVENTION: Thermostable luciferase Of Firefly,
; TITLE OF INVENTION: Thermostable luciferase Gene Of Firefly, No. 5229285el Reco
; TITLE OF INVENTION: DNA, And Process For The Preparation Of Thermostable
; TITLE OF INVENTION: Luciferase Of Firefly
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennile & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/903,047
; FILING DATE: 19920623
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7005-048
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1644 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA to mRNA
; US-07-903-047-1
alignment_scores:
Quality: 2684.00 Length: 548
Ratio: 4.943 Gaps: 0
Percent Similarity: 99.088 Percent Identity: 93.431
alignment_block:
US-09-581-241-6 x US-07-903-047-1 ..
Align seg 1/1 to: US-07-903-047-1 from: 1 to: 1644
1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
|||||
1 ATGGAATAATGATGAGAAATATTTGTACTGACCTTAACCGTT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMet 34
|||||
51 TTACCTATATGAGAGGAGATCTGCGAACACAAATTAAGCAATACATG 100

```

```

1001 ATCTTCCCGGTGTCSTCAAGGTATGGTTTAACGAACAACATCTGCC 1050
351 TlelllellethrProglunglyasprlsyProglualaserglylsva 367
|||||
1051 ATTATTATTACACCAGAAGGACGATTAACCGAGGCTTGGAAGAAGT 1100
367 ValProleuphelysAlalysValleaspleuasprhlrlyslysthrL 384
|||||
1101 CGAGCGGTGTTTAAAGCAAAGTATTGATCTTGATACAAAAAATCTT 1150
384 euclgProsnatgatvgclgyduvalcylsvAllysGlyPrometleumet 400
|||||
1151 TAGCTCCTAACACAGCTGGAGAAGTTTGTTTAAAGCACTCATCTTAATG 1200
401 LysGlYTYrValAsprasnProglunlatThrArgluilelleaspglugl 417
|||||
1201 AAGGTTATGTAAATNATATCCAGACAGCAAAAAGACTTATGGAGAAGA 1250
417 uglYTrPLeuhIsthrglyAsplleGlyTyrrTyraSpGluuglyshisp 434
1251 AGGTTGCTGCACACCGSAGATATTGGATTATATGATAGAAAAACATT 1300
434 herhelllevalasparglyserleullelytyrlysglytyrciln 450
1301 TCCTTATTGTGCATCGTTGGAAGTCTTTATCAATAACAAAGATACCAA 1350
451 ValProfoalagluleugluservalleuleuglnhisproasnlleph 467
1351 GTACACCTCCGCAATTAGAAATCCGTTCTTTGCAACATCCATCATCTT 1400
467 eaSprAlglyValAlaglyValIProasprroilleaglygltuleuprog 484
1401 TGATGCGGGGTGGCCGCGCTTCGATCCTGTAGCTGGCGAGCTTCAG 1450
484 lYAlaValAlvalleulylsGlylsGlySermetThrgluLygslyuval 500
1451 GAGCGGTGTTGTACTGGAAGCGGAAAAAATATGACCCGAAAAAGAGTA 1500
501 MetaspYrYValAlaserGlyValSeranAlalyArtyrleuArGlygl 517
1501 ATGATATTATGTTCMAAGTCAAGTTCAAAAGCAAAACGTTTACGTGGTGG 1550
517 yValArpHevalaspslyuvalProlysglyeuthrglylyslleaspg 534
1551 TGTCGTTTTTGSATCAAGTACTTAAGGCTTACTGGAAAAATTGGAGG 1600
534 lyLysAlAlleArtyrGluileuleulylsProValAlalysMet 548
1601 GCAGAGCAATTAGAGAAATCCTTAAGAAACCACTTGCTTAAGATG 1644

seq_name: /cogn2_6/prodata/2/ina/5A_COMB.seq;US-08-076-042-1

seq_documentation_block:
Sequence 1, Application US/08076042
Patent No. 5330906
GENERAL INFORMATION:
APPLICANT: KAJIYAMA, NAOKI
APPLICANT: NAKANO, EIICHI
TITLE OF INVENTION: MUTANT LOCIFERASE OF A FIREFLY, MUTANT
TITLE OF INVENTION: LOCIFERASE GENES, NOVEL RECOMBINANT DNAS CONTAINING THE
TITLE OF INVENTION: GENES AND A METHOD OF PRODUCING MUTANT LOCIFERASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: N.Y.
COUNTRY: U.S.A
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/076,042
 FILING DATE: 15-JUN-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/675,211
 FILING DATE: 26-MAR-1991
 NAME: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MISROCK, S. LESLIE
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7005-026-999
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1644 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 ORIGINAL SOURCE:
 ORGANISM: *Unciola cruciata*
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1644
 US-08-076-042-1

alignment_scores:
 Quality: 2684.00 Length: 548
 Ratio: 4.943 Gaps: 0
 Percent Similarity: 99.088 Percent Identity: 93.431

alignment_block:
 US-09-581-241-6 x US-08-076-042-1 ..

Align seg 1/1 to: US-08-076-042-1 from: 1 to: 1644

```

1 MetGuaaMeGluAsnAspGluAsnIleValIYrGlyProGluProPh 17
1 ATGGAAACATGGAAGAAACGATGGAATTTGTTGAGTGAACCTTAACCGTT 50
17 eTyPProIleGluGluGlySerAlaGlyAlaGlnIleuArgLysTyrMetA 34
51 TTACCCATTCGAGAGGAGATCTGCTGACACACATTTACGCAATACATGG 100
34 sPArGTYrAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 50
101 AGCGATATGCAAAACCTTGCGCAATTCCTTACAAATGCAAGTACTGCTGT 150
51 ValAspTYrThrTYrAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 67
151 GTTATATTCTTACGCCGAATACCTTGGAAGAAATCATGTTGTTAGGAAA 200
67 uAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 84
201 AGCTTTGCAAAATATTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 250
84 eArgLysGlyGluGluPhePheIleProValIleuAlAlAlAlAlAlAlAl 100
251 GTGAAACCTGGAAGAAATTTTATCTCTTAATAGCCGAGACTGTTATA 300
101 GLYValGlyValAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 117
301 GGTGTAGAGTGTGCAACCACTAATGAGATTACCTTACCTGTAACCTGCT 350
117 lHsSerLeuGlyIleSerLysProThrIleValIlePheSerLysLysG 134

```

```

351 TCACAGTTTAGGTATCTCTTAACCAACATTTGATTAGTTCTTAACAAAG 400
134 lYleuAspLysValIleThrValGlnLysThrValThrAlAlAlAlAlAl 150
401 GCTTACATAAAGTTATACAGTACAGAAACAGTAACACTACTATTAAACC 450
151 lIleValIleleuAspSerLysValAspTYrArgLYrGlnSerMetAs 167
451 ATTGTATATCTAGATAGCAAAAGTTGATTATCGAGATATCAATGCTCGA 500
167 pAsnPhelIleLysAsnThrProGlnIlePheLysGlySerSerPheI 184
501 CACCTTATATAAAGAAAGAACTCCACAGGTTTCAAGATCCAGTTTCA 550
184 ySthrValGluValAsnArgLysGluGlnValAlAlAlAlAlAlAlAlAl 200
551 AAACGTGGAAGTTGACCGTAAAGAAAGAGTTCCTTATATGAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnIleuThrHisGlnAsn 217
601 TCGGTTCTACCGGTTTGCAGAAAGCGTACAACTTACACGAAATATAC 650
217 eValThrArgPheSerHisAlaArgAspProIleTYrGlyAsnGlnVal 234
651 AGTCACTAGATTTTCTCATGCTAGATCCGATTTATGTTAGTAAACAGTT 700
234 ePrGlyThrAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 250
701 CACCAAGCACCGCTGTTTAACTGTCATCCATTCATCATGTTTGGT 750
251 MetPheThrThrLeuGlyTYrLeuThrCysGlyPheArgIleValIle 267
751 ATGTCACACTACTAGGATATTATTTGTTGTTGTTGTTGTTGTTGTTG 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTYrLysC 284
801 AACAAATTCGATGAAGAAACATTTTAAACCTACAAAGATTAAAT 850
284 ySerSerValIleleuValProThrIleuPheAlAlIleuAsnArgSer 300
851 GTTCAAGTTTATCTTGTACCGACCTGTTTGCATTCCTCAACAAAGT 900
301 GluLeuLeuAspLysTYrAspLeuSerAsnLeuValGluIleAlAlSerG 317
901 GAATATCAATTAATATACGATTTGTCAATTTAGTTGATGATGATCTGG 950
317 yGlyAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 334
951 CGGACACCTTTATCAAAAGAAAGTTGGAGCTGTTGCTAGACCTTTTA 1000
334 snLeuProGlyValAlArgGlnIleTYrGlyLeuThrGluThrThrSerAl 350
1001 ATCTCCCGGCTGTTGTCAGGTTATGTTTACAGAAACAAACATCTCC 1050
351 lIleIleIleThrProGluGlyAspAspLysProGlyAlAlAlAlAlAl 367
1051 ATTATATTACACAGAGAGAGATTAACACAGAGCTTCGGAAGAGT 1100
367 lValProLeuPheLysAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 384
1101 CGTGCGCTGTTTAAAGCAAAAGTTATGATCTGATCCAAAAAATCTT 1150
384 eArgLysProAsnArgArgLysGluValCysValLysGlyProMetLeuMet 400
1151 TAGGCTTAACAGAGAGTGAAGATTTGTTTAAAGCACTATGCTTATG 1200
401 LysGlyTYrValAlAspAsnProGluAlAlAlAlAlAlAlAlAlAlAl 417
1201 AAAGTTATCTTAATTAATCCAGAAACAAACAAAGCTTATGACGAAGA 1250
417 uGlyTYrLeuHisThrGlyAspIleGlyTYrTYrAspGlnLysHisP 434
1251 AGTTTGCTGCACACCGAGATATTGATTTATGATGAAGAAAAACATT 1300

```



```

434 hepellevalaspargleuylserleuileuyltyrlysglytyrln 450
|||||
1301 TCTTATTGTCGATCGTTTGAAAGCTTTATCAAAATCAAGGATACCAA 1350
|||||
451 ValProprolactinleuylserleuileuylhisprosnleph 467
|||||
1351 GTACCCACTCCGAAATAGATCCGTTCTTTGCAACATCCATCTATCTT 1400
|||||
467 easpalaglvalalaglvalproaspproilealagluleuprog 484
|||||
1401 TGATGCGGTGGCGCGCTTCCTGATCTGTAGCTGGCGACTTCGAG 1450
|||||
484 lvalalvalalvalleuylslyglysermethrglylulysgluval 500
|||||
1451 GACCCGTTGTGTACTGGAAGCGGAAAAATATGACCGAAAAAAGATA 1500
|||||
501 Metaspyrvalalaserlnvalaserlnalalysargleuarglyl 517
|||||
1501 ATGGATTATGTCGACGTCAGTTTCAAAATGCAAAAGCTTTAGCTGGTGG 1550
|||||
517 yvalargphevalaspcluvalprolysglyleuthrglyllysleapg 534
|||||
1551 TGTGCTGTTTGTGATGAGTAAAGCTTACTGAAAAATGACG 1600
|||||
534 lylsalaileargluileuylslysprovalalalysmet 548
|||||
1601 GCACAGCAATTAGAGAAATCCTTAAGAACCAAGTGTCTAAGATG 1644
|||||
seq_name: /cgn2_6/prodata/2/lna/5B_COMB.seq:US-08-757-046A-3

```

```

seq_documentation_block:
; Sequence 3, Application US/08757046A
; Patent No. 5876995
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; TITLE OF INVENTION: BIOLUMINESCENT ARTICLES OF MANUFACTURE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,046A
; FILING DATE: 11-25-96
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/597,274
; FILING DATE: 02-06-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6680-105B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; TELER:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1644 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:

```

```

; NAME/KEY: Coding Sequence
; LOCATION: 1..1644
; OTHER INFORMATION: Luciola Crucifera Luciferase (Firefly)
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: PATENT NO.: 4,968,613
; US-08-757-046A-3

alignment_scores:
  Quality: 2684.00      Length: 548
  Ratio: 4.943          Gaps: 0
  Percent Similarity: 99.088  Percent Identity: 93.431

alignment_block:
US-09-581-241-6 x US-08-757-046A-3 ..

Align seg 1/1 to: US-08-757-046A-3 from: 1 to: 1644

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProH 17
|||||
1 ATGGAAGACATGGAAGACATGGAAGATATTGATGACCTAAACCGTT 50
|||||
17 eTyrProIleGluGlySerAlaGlyAlaGluLeuArgLysTyrMetA 34
|||||
51 TTACCTATCGAAGAGGATCTGCTGGAAACAAATTACGCAAAATCATG 100
|||||
34 spArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
|||||
101 AGCGATATGCAAAACTGGCGCAATTGCTTTACAAATCAGTACTGCT 150
|||||
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLysLeuGly 67
|||||
151 GTTGATATTATCTTACCGCGAATCTTGAGAAATCATGTTGCTAGAAA 200
|||||
67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuCys 84
|||||
201 AGCTTGCAGAAATATGATGTTGTTGTTGATGGCAAAATGGCTTAGCA 250
|||||
84 eArgLysCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
|||||
251 GTGAAACCTGGAAGATTTTATTCCGTATATACCGGACGTTTATA 300
|||||
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
|||||
301 GGTTAGTGTGTCACCCACTAATGAGATTACACTTACGTAACGTGGT 350
|||||
117 HisSerLeuGlyIleSerLysProThrIleValIlePheSerLysLysG 134
|||||
351 TCACAGTTTAGTATCTCTAACCACCAATTTGATTTAGTTCTTAAAGAG 400
|||||
134 LysLeuAspLysValIleThrValGluLysThrValThraIleLysThr 150
|||||
401 GCTTAGATAAGTATATACAGTACAGAAACAGTACTATTAAAC 450
|||||
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
|||||
451 ATTGTTACTAGATACCAAGTTGATTATCGAGAAATCAAGTCTGGA 500
|||||
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerPheL 184
|||||
501 CACCTTTATTAAGAAAGAACACTCCACAGGTTTCAAGCATCCAGTTTCA 550
|||||
184 yThrValGluValAsnArgLysGluGluValAlaLeuIleMetAsnSer 200
|||||
551 AAACGTGGAAGTGTGACCGTAAAGACAAGTGTCTTATTAAGAACTCT 600
|||||
201 SerGlySerThrGlyLeuProLysGlyValGluLeuThrHisGluAsnI 217
|||||
601 TCGGTTCTTACCGGTTTGCCAAAGGCGTACAACTTACTACGAAATAC 650
|||||
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGluValS 234
|||||
651 AGTCACTAGATTTTCTCATGCTAGAGATCCGATTATGTAGCAACAATT 700
|||||

```

```

234  exProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
|||||
701  CACGAGCAGCAGCGCTGTTTAACGTGCTTCATTCATCATGTTTGGT 750
251  MetPheThrThrLeuGlyTyThrLeuThrCysGlyPheArgIleValMetIle 267
|||||
751  ATGTTCTACACTCTAGGGAATTAATTTGTTGTTTCTGTTGTTAAATGTT 800
267  uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
|||||
801  AACAAATTCGATGAGAGAAACATTTTAAAACTTACAAAGATTATTAAT 850
284  ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
|||||
851  GTACAAAGTGTATTTCTTTGACGACCTGTTTGCATTTCTCAACAAAGT 900
301  GluLeuLeuAspLysTyrAspLysSerAsnLeuValGluIleAlaSerG 317
|||||
901  GAATTAATCTCAATTAATACGATTGTTCAAAATTAATTAATGAGATTGCA 950
317  yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
|||||
951  CGAGCAGACCTTATCAAAAGAGTTGCTGAGCTGTTGCTAGACGCTTTA 1000
334  snLeuProGlyValAlaArgGlnGlyTyThrGlyLeuThrGluThrSerAla 350
|||||
1001  ATCTTCCCGGTGCTGTCAGAGTTATGTTTAAACAGAAACAACTGCGC 1050
351  IleIleIleThrProGluGluLysAspLysProGlyAlaSerGlyLysVa 367
|||||
1051  ATTATTATTACACAGAGAGACATTAACCAAGAGCTTCTGAGAAAGT 1100
367  lValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
|||||
1101  CGTGCGCTGTTTAAAGCAAAAGTATTGATCTTGATACCAAAATCTT 1150
384  euGlyProAsnArgArgGlyGluValLysValLysGlyProMetLeuMet 400
|||||
1151  TAGGTCCTTAACGAGCTGAGAGAGTTGTTGTTAAAGACCTAATGCTTAA 1200
401  LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
|||||
1201  AAAGTATATGAATAATCAACAGAACCAAAAGAACTTATTGAGAGAA 1250
417  uGlyTyrLeuHisThrArgLysAspIleGlyTyrTyrAspGluGluLys 434
|||||
1251  AGGTTGCTGACACCGGAGATATTGATATGATGAGAGAAAAACATTT 1300
434  hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
|||||
1301  TCTTATTATGTCGATCGTTTGAAGTCTTTAATCAAAATACAAAGGATAC 1350
451  ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIleP 467
|||||
1351  GTACCACCTGCGCAATTAATCCGTTCTTTTGCACATCATCATCTACTT 1400
467  eaSPAlaGlyAlaIleGlyValProAspProIleAlaGlyGluLeuProG 484
|||||
1401  TGAATGCTGGTGTTCGCGGCTTCTGATCTGCTGCGAGCTTCCAG 1450
484  lYalAlaValAlaValLeuLysLysGlyLysSerMetThrGluLysGlu 500
|||||
1451  GAGCGGTTGTTGCTGAGAAAGCGGAAAAAATATGACCGGAAAAAGAACT 1500
501  MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
|||||
1501  ATGATATTATGTCAGACAGTTCAAATCAAAACGTTTACGCGTGG 1550
517  yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
|||||
1551  TGTTGTTTTGTGGATGAAGTACTAAAGGTCTTACTGGAATAATGACG 1600

```

```

534  lYlSaAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
|||||
1601  GCAGAGCAATTAAGAAATCTTAAAGAAACCACTGCTGAAGATG 1644
seq_name: /cgn2-6/plodata/2/ina/6A_COMB.seq:us-09-447-208-3

```

```
seq_documentation_block:
```

```
; Sequence 3, Application US/09447208
```

```
; Patent No. 6113886
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Bryan, Bruce
```

```
; TITLE OF INVENTION: BIOLUMINESCENT ARTICLES OF MANUFACTURE
```

```
; NUMBER OF SEQUENCES: 14
```

```
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Heller Ehrman White & McGrath
```

```
; STREET: 4250 Executive Square, 7th Floor
```

```
; CITY: La Jolla
```

```
; STATE: CA
```

```
; COUNTRY: USA
```

```
; ZIP: 92037
```

```
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Diskette
```

```
; COMPUTER: IBM Compatible
```

```
; OPERATING SYSTEM: DOS
```

```
; SOFTWARE: FastSeq Version 1.5
```

```
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/09/447,208
```

```
; FILING DATE:
```

```
; CLASSIFICATION:
```

```
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: 0909/135,988
```

```
; FILING DATE: 08-17-98
```

```
; CLASSIFICATION:
```

```
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: 08/757,046
```

```
; FILING DATE: 11-25-96
```

```
; CLASSIFICATION:
```

```
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: 08/597,274
```

```
; FILING DATE: 02-06-96
```

```
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Seidman, Stephanie L
```

```
; REGISTRATION NUMBER: 33,779
```

```
; REFERENCE/DOCKET NUMBER: 24727-105C
```

```
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 619-450-8400
```

```
; TELEFAX: 619-450-8499
```

```
; TEXT:
```

```
; INFORMATION FOR SEQ ID NO: 3:
```

```
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 1644 base pairs
```

```
; TYPE: nucleic acid
```

```
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: cDNA
```

```
; FEATURE:
```

```
; NAME/KEY: Coding Sequence
```

```
; LOCATION: 1...1644
```

```
; OTHER INFORMATION: Luciola cruciata Luciferase (firefly)
```

```
; PUBLICATION INFORMATION:
```

```
; US-09-447-208-3
```

```
alignment_scores:
```

```
Quality: 2684.00 Length: 548
```

```
Ratio: 4.943 Gaps: 0
```

```
Percent Similarity: 99.088 Percent Identity: 93.431
```

```
alignment_block:
```

```
US-09-581-241-6 x US-09-447-208-3 ..
```

```
Align seg 1/1 to: US-09-447-208-3 from: 1 to: 1644
```

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
1 ATGGAAGAAACATGAGAAACATGAGAAATATGTAGTGGACCTAAACCGTT 50
17 eTyrProIleGluGlySerIleGlyIleGluIleArgLysTyrMetA 34
51 TTAACCCATATGAGAGAGGATCTCTGGACACATTAACGCAATACATGG 100
34 sPArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 AGCGATATGCAAAACTTGGCGCAATGCTTTACAATGCAAGTTACTGCT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
151 GTTGTATATCTTACGCCGATCTGGAGAAATCAATGTTGTCTAGGAAA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
201 AGCTTTGCCAAATATGTGTTGGTTGGTTGATGGCAGAAATGCGTTATGCA 250
84 eArgLysAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
251 GTGAAGAACTGGAAGAAATTTTATCTCTGTAATAGCGCGACTGTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
301 GGTGTAGGTGTGGACCCACCAATAGATTTACCTTACCTGACGTAAGTGT 350
117 LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
351 TCACAGTTTATGATCTCTTAACCAACAATGTTAGTTAGTTCTAAAAAG 400
134 LysLeuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150
401 GCTTGTAGTAAAGTATATACACTACAGAAACAGTAACCTACTATTAAAAAC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
451 ATTGTATATCTAGATAGCAAGATGATATGAGATATCAATGCTGTGCA 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerPheIle 184
501 CACCTTTATAAAAAGAAACACTCCACAGGTTTCAACGATCCAGTTTCA 550
184 YSThrValGluValAsnArgLysGluGluValAlaIleLeuIleMetAsnSer 200
551 AAACGTGTGAAGTGTACCGTAAAGAACAAAGTGTCTTATATGTAAGTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGluLeuThrHisGluAsnIle 217
601 TCGGGTTCTACCGGTTTGCAGAAAAGCGTACAACTTACTACAGAAAATAC 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGluValS 234
651 AGTCACTAGATTTTCTCATGCTAGAGATCCGATTTATGTGTAAACAAAGTT 700
234 eArgProGlyThrAlaIleLeuThrValValProPheHisGlyPheGly 250
701 CACCAAGCACCGCGTGTATTAACGTGCTTCATTCATCATGTTGTGTGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
751 AAGTTCTACACTCTACGATATTTAATTTGTGTGTTTCTGTTGATATCT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGluAspTyrLysC 284
801 AACAAATTCGATGAAGAAACATTTTAAAACTCTACAGATTAATAAT 850
284 YSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTACAACTGTTATCTGTGTAACGACCTGTGTTCGATTCGCAACAAAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317

```

```

901 GAATTAATCAATAAATACGATTTGTCAAAATTTAGTGAATTCGATCTGG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
951 CGAGCACCTTATCAAAAAGAGTGTGAACCTGCTGCTAGACCTTTA 1000
334 sPLeuProGlyValAlaArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
1001 ATCTTCGCGGTGTGTCTCAAGTTATGTGTTTAAAGAAACAAACATCTCC 1050
351 IleIleIleThrProGluLysAspLysProGlyAlaSerGlyLysVal 367
1051 ATTATATTATCAACAGAAAGAGACGTAACACAGAGCTTCTGAAAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
1101 CGTGCCGTTGTTTAAAGCAAAAGTATGTATCTGTATACCAAAAATCTT 1150
384 eGluGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
1151 TAGGCTCTACAGACGTGGAGAAATTTGTTAAAGACCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
1201 AAAGTTATCTAATAATATCCAGAACACAAAGAACTTATGACGACAGA 1250
417 uGlyTyrPheHisThrGlyAspIleGlyTyrTyrAspArgGluLysHisP 434
1251 AGGTGGCTGCACCGCGAATATGTGATTTATGTATGTAAGAAAAACAT 1300
434 hPheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTTATTTGCGATCGTTTGAAGCTTTTATCAAAATACAAAGGATCCAA 1350
451 ValProProAlaGluLeuGluLysValIleLeuGluHisProAsnIlePh 467
1351 GTACCACTCCCGCAATTAGAATCCGTTCTTTCACACATCAATCTATCT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLysLeuProG 484
1401 TGATGCGGTGTGTCGCGGCTCTCTGATCTCTGATGCGGAGACTTCCAG 1450
484 lValAlaValAlaLeuLysLysGlyLysSerMetThrGlyLysGluVal 500
1451 GAGCGGTGTGTGTACTGGAAGCGGAAAAAATATGACCGCAAAAAAGTGA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
1501 ATGGATTAATGTTGCAAGTCAAGTTTCAAAATGCAAAACGTTTACGTGTG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
1551 TGTCGCTTTGTGTGATGAAGTAAAGTCTTACTAGGAAAAATGAGC 1600
534 lLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
1601 GCAGAGCAATTAGAGAAATCTTAAAGAACCGATGTTGCTAAGATG 1644
seq_name: /cgn2_6/prodata/2/lna/6A_COMB.seq:US-09-135-988-3
seq_documentation_block:
: Sequence 3, Application US/09135988
: Patent No. 6152358
: GENERAL INFORMATION:
: APPLICANT: Bryan, Bruce
: TITLE OF INVENTION: BIOluminescent ARTICLES OF MANUFACTURE
: NUMBER OF SPOUNCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Heller Ehrman White & McCaulliffe
: STREET: 4250 Executive Square, 7th Floor
: CITY: La Jolla
: STATE: CA

```

```

? COUNTRY: USA
? ZIP: 92037
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FastSeq Version 1.5
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/135,988
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/757,046
? FILING DATE: 11-25-96
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/597,274
? FILING DATE: 02-06-96
? ATTORNEY/AGENT INFORMATION:
? NAME: Seidman, Stephanie L
? REGISTRATION NUMBER: 33,779
? REFERENCE/DOCKET NUMBER: 24727-105C
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619-450-8400
? TELEFAX: 619-450-8499
? TELEX:
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1644 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: Coding Sequence
? LOCATION: 1..1644
? OTHER INFORMATION: Lucifolia Crucifera Luciferase (Firefly)
? PUBLICATION INFORMATION:
? US-09-135-988-3

alignment_scores:
  Quality: 2684.00      Length: 548
  Ratio: 4.943          Gaps: 0
  Percent Similarity: 99.088      Percent Identity: 93.431

alignment_block:
  US-09-581-241-6 x US-09-135-988-3 ..

Align seg 1/1 to: US-09-135-988-3 from: 1 to: 1644

```

```

251 GTGAAACTGTGAAGATTTTTCCTGTAATAGCCGACGTGTATA 300
101 GYVALGlyValAlaProThrAsnGluIleTyThrLeuArgGluLeuVa 117
301 GGCTAGAGTGTTCACCCACTAATGAGATTACACTTACGTGACACTG 350
117 IHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
351 TCACAGTTTAGTATGCTTAAACCAACAAATGTAATTTAGTTCTAA 400
134 ILeuAspLysValIleThrValGlnSerThrValIleAlaIleY 150
401 GCTTAGATAAAGTTATTAACAGTACAGAAACAGTAATCTATTA 450
451 ATTGTATACAGATGACCAAGTTGATTCGAGATATCAATGCTGCA 500
167 PAsnPhelIleLysLysAsnThrProGlnGlyPheLysGlySer 184
501 CACCTTTATTAAGTAAGAACACTCCACAGGTTTCAAGCAAGTTT 550
184 YThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAs 200
551 AAAGTGTGAAGTTGACCGTAAGAACAAAGTCTCTTATTAAGACT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisLu 217
601 TCGGTTTCAACGGTGTTCGCAAAAGCGTAACAATCAACCAAAAT 650
217 eValThrArgPheSerHisAlaArgAspProIleTyArgLysG 234
651 AGTACAGATTTTTCATGCTAGATCGATTCGATTCGATTCGAT 700
234 eProGlyThrAlaIleLeuThrValValProPheHisGlyPhe 750
701 CACGAGCAGCGCTGTTTAACTGCTGCAATCCATCATCATGAT 750
251 MetPheThrThrLeuGlyTyThrLeuThrCysGlyPheArgI 267
751 ATGTCACACTACTAGGTATTTAATTTGTTGTTGTTGTTGTTG 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAsp 284
801 AACAAATTCGATGAAGAAACATTTTAAAACTACAAAGATTA 850
284 YSerSerValIleLeuValProThrLeuPheAlaIleLeuAs 300
851 GTACAGGTATTTCTGTACCGACCTGTTGCAATTCCTCAAA 900
301 GluLeuLeuAspLysTyAspLeuSerAsnLeuValGluIleAl 317
901 GAATTCCTCAATAATATACGATTTGTCAAATTTAGTTGAGTT 950
317 YGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArg 334
951 CGAGACACTTATCAAAAGAGTGGTGAACCTGTTCTGACCT 1000
334 sLeuProGlyValArgGlnGlyTyThrGlyLeuThrGluThr 350
1001 ATCTCCCGCGTGTCTGACAGTTATGGTTTAAACGAAACAC 1050
351 IleIleIleThrProGlnGlyAspAspLysProGlyAlaSer 367
1051 ATTATTATTACACGAGAGAGACGATTAACGAGAGCTTCG 1100
367 IValProLeuPheLysAlaLysValIleAspLeuAspThrL 384
1101 CGTCCGCTGTTGTTAAAGCAAAAGTTATTGATCTGATACCA 1150
384 eGluProAsnArgArgLysGluValCysValLysGlyProMet 400
1151 TAGGCTTAACAGAGCTGAGAGAGTTGTTTAAAGGACCTAT 1200

```

```

401 LysGlyTyrValAspAsnProGluAlaThrArgGluLeuIleAspGlu 417
|||||.....
1201 AAGGTATGTAATATATCCGAAAGCAACAAAGAACTTATTGACGAGA 1250
417 uGlyTyrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluHisP 434
|||||.....
1251 AGTTGGCTGCACACCGAGATATGATATTTATGATGAAGAAAACATT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
|||||.....
1301 TCTTATTTGCGATCGTTGAGTCTTATTAACAATGCAAGATACCAA 1350
451 ValProPheAlaGluLeuGluSerValLeuLeuGlnHisProAsnIleP 467
|||||.....
1351 GTACCACCTGCGCAATTAGAAATCCGTTCTTTGCAACATCATCTATCT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
|||||.....
1401 TGATGCTGCTGTTCCCGCGCTTCTGATCTGTAGCTGGCGAGTTCCAG 1450
484 LysAlaValAlaValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
|||||.....
1451 GAGCGGTGTTGTACTGAAAGCGGAAATAATATGACCGAAAGAAAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
|||||.....
1501 ATGATTTATGTTGCAAGTCAAGTTTCAATGCAAAAGCTTACGTGTGG 1550
517 ValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
|||||.....
1551 TGTTGCTTTTGTGATGAGTAAAGTCTTACTGGAATAAATTCAGC 1600
534 LysLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
|||||.....
1601 GCAGAGCAATTAGAGAAATCCTTAAGAAACACAGTCTTAAGATG 1644
seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:us-09-277-716-3
seq_documentation_block:
; Sequence 3, Application US/09277716A
; Patent No. 6232107
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
; CURRENT APPLICATION NUMBER: US/09/277,716A
; EARLIER FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 60/102,939
; EARLIER FILING DATE: 1998-10-01
; EARLIER APPLICATION NUMBER: 60/089,367
; EARLIER FILING DATE: 1998-06-15
; EARLIER APPLICATION NUMBER: 60/079,624
; EARLIER FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1644
; TYPE: DNA
; ORGANISM: Luciola cruciata
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1644)
; FEATURE:
; OTHER INFORMATION: Luciola cruciata (firefly) luciferase
; PUBLICATION INFORMATION:
; PATENT DOCUMENT NUMBER: 4,968,613
; PATENT FILING DATE: 1988-07-26
; PUBLICATION DATE: 1990-11-06
; US-09-277-716-3

```

alignment_scores:

Quality: 2684.00 Length: 548
Ratio: 4.943 Gaps: 0
Percent Similarity: 99.088 Percent Identity: 93.431

alignment_block:

US-09-581-241-6 x US-09-277-716-3 ..

Align seg 1/1 to: US-09-277-716-3 from: 1 to: 1644

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProP 17
|||||.....
1 ATGGAACACATGGAACGATGAAATATTTGATGTTGGACCTTAACCGTT 50
17 eryTyrProIleGluGlySerAlaGlyValGlnLeuArgLysTyrMet 34
|||||.....
51 TTACCTATGCAAGAGAGATCTGCTGGAACACAAATTACCAATATCATG 100
34 sParGTYrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
|||||.....
101 AGCGATATGCAAAACTGGCGCAATTGCTTTACAATGCAATGCTACTGCT 150
51 ValAspTyrThrTyrAlaGlyTyrLeuGluLysSerCysCysLeuGly 67
|||||.....
151 GTTGATTATTCCTTACCGCGCAATTAATGGAATATCATGTTGCTAGGAA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgGlyIleAlaLeuCys 84
|||||.....
201 AGCTTTCGAAATTAATGATGTTGTTGTTGATGACCAATTCGTTATGCA 250
84 erGluAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
|||||.....
251 GTCAAAACGTGCAAGAAATTTTATTCCTGTAATATACCGGACGTTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
|||||.....
301 GGTGTAGCTGTTGCACCCACTAATGAGATTACACTTTACGTAACTCGCT 350
117 LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
|||||.....
351 TCACAGTTTAGTATCTCTTAACCAACAAATTGATTTAGTTCAAAAAG 400
134 LysLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
|||||.....
401 GCTTAGATAAAGTTATTAACAGTACAGAAACACGTAATCTATTAACCC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMet 167
|||||.....
451 ATGTTTATTAAGTATGCAAGATGATTAATGAGATATCAATGCTGGA 500
167 pAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
|||||.....
501 CACCTTTATATAAAGAAACACCTCCACGAGTTTCAAGCATTCAGTTTCA 550
184 yStrThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
|||||.....
551 AAACGTGGAAGTTGACCGTAACACAGATGCTCTTAATAAGAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnI 217
|||||.....
601 TCGGCTTCTACCGGTTTCCCAAAAGCGGTACAATCTCACGAAATATAC 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
|||||.....
651 AGTCACTAGATTTCTCATGCTAGAGATCCGATTTATGTTAACCAGATT 700
234 erProGlyThrAlaIleLeuThrValAlaProPheHisHisGlyPheGly 250
|||||.....
701 CACACGCAACCGCTGTTTAACGTCGTTCCATTCATCATGATGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMet 267
|||||.....
751 ATGTTCACTACTAGAGTATTAATTGATGCTGTTGCTGTTGAATGTT 800

```

```

267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
801 AACAAATTCGATGAGAAACATTTTAAAACTCTACAGATTATAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuSnaArgSer 300
851 GTACAGAGTTATCTGTACGACCTTGTTCATCTCTCAACAAACAT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
901 GAATTACTCAATAATACGATTGTTCAAAATTAGTTGACATTGCAATCTG 950
317 yGlyAlaProLeuSerLysGluIleGlyLuuAlaValAlaArgArgPhea 334
951 CGAGACCTTTATCAAAAGAGTTGTAAGCTGTTCTGACAGCTTTA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAla 350
1001 ATCTTCCGGTGTGTCAGGTTATGTTTAAACGAAACCAATCTGCC 1050
351 IleIleIleThrProGlnGlyAspAspLysProGlyAlaSerGlyLysVa 367
1051 ATTATATTATACACGAGAACGATTAACACGAGCTTCTGAAAACT 1100
367 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysTyrL 384
1101 CGTGGCGTGTTTAAAGCAAAAGTTATGATCTGTATACCAAAATCTT 1150
384 euGlyProAsnArgArgGlyGlyValGlyValGlyLysGlyProMetLeuMet 400
1151 TAGGTCCTACAGACGTGAGAGAGTTGTGTTAAAGACCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleLeuAspGluG 417
1201 AAAGGTATATGTAATATACCAAGACCAACAAAGACTTATGACAGAGA 1250
417 uGlyTrpLeuHisThrGlyAspIleGlyTyrTyrAspGluGlyHisP 434
1251 AGCTTGCGTGCACCGCAGATTTGATATGATGATAAAGAAACATTT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTATATGTCGATGCTTGAAGTCTTTAATCAATTAACAAAGATACCA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
1351 GTACACCTGCCGAATTAGAATCCGTTCTTTCGACACATCCATCTATCTT 1400
467 eaSpAlaGlyValAlaGlyValProAspProIleAlaGlyLuuLeuProG 484
1401 TGATGCTGTGTTGCCGCGCTTCTGATCTGTTAGCTGCGAGCTTTCAG 1450
484 lAlaValAlaValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
1451 GAGCGCTGTGTTGTAAGCGGAAAGAAATATGACCGCAAAAGAGCTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG 517
1501 ATGGATTTATGTTGCAGATCAAGTTCAAAATGCCAAAGCTTACGTTG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysTyrL 534
1551 TGTTCGTTTGTGATGATGATCTTAAGGCTTACTGAGAAAATTTACG 1600
534 lYlyAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
1601 GCAGAGCAATTAAGAAATCTTAAAGAACACAGTTGCTAAGATG 1644
seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-08-597-274A-3
seq_documentation_block:
; Sequence 3, Application US/08597274A

```

```

; Patent No. 6247995
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; TITLE OF INVENTION: BIOLUMINESCENT NOVELTY ITEMS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/597,274A
; FILING DATE: 02/06/96
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6680-105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1644 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1..1644
; OTHER INFORMATION: Luciola cruciata Luciferase (firefly)
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: 4,968,613
; US-08-597-274A-3

alignment_scores:
    Quality: 2684.00      Length: 548
    Ratio: 4.943          Gaps: 0
    Percent Similarity: 99.088      Percent Identity: 93.431

alignment_block:
US-09-581-241-6 x US-08-597-274A-3 ..
Align seg 1/1 to: US-08-597-274A-3 from: 1 to: 1644

1 MetGlnAsnMetGluLysAspGluLysAsnIleValTyrGlyProGluProPh 17
1 ATGGAAACATGTGAAAGACGATGAAATATTTAGTTGACCTTAAACCGTT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetA 34
51 TTACCTATCGAAGAGGATCTGTGGAACACAAATTACGCAATACATG 100
34 sPARgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 AGCGATATGCAAAACCTGGCGCAATGCTTTACAAATGCAATGCTACTG 150
51 ValAspTyrThrTyrAlaGluThrLeuGlnLysSerGlyLysGlyGluG 67
151 GTTGATTTATCTTACGCCGCAATCTTGGAGAAATCATGTTGTTAGGAA 200

```

```

67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCysS 84
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 AGCTTTCAGAAATTAATGTTGGTTGGTTGATGGCAGAAATTCGTTAAGCA 250
84 ergLysAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
251 GTGAAACCTGTGAAGAAATTTTATTCCTTAATAGCGGCACTGTTATA 300
101 GlYValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
301 GGTTAGAGTGTGGCACCCTAATAGACATTACACTTACGTGAACGTGGT 350
117 LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLysG 134
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
351 TCACAGTTAGTATCTCTAAACCAACAATTTGATTTGTTCAAAAAG 400
134 LysLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
401 GCTTAGATAAAGTTATACAGCTACAGAAACAGTAACACTATTAAACC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
451 ATGTGTAATAGATAGCAAAAGTTGATTATCGAGATATCATGCTCGGA 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheL 184
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
501 CACCTTTATTAAGAAACACTCCACCAAGGTTTCAAGCATCCAGTTTCA 550
184 YStrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsSer 200
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
551 AAACGTGGAAGTTGACCGTAAACAACAAGTTGCTTATATATCAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnI 217
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
601 TCGGGTCTACCGGTTGGCAAAAGGCGTACACTACTCAGCAAAATATAC 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
651 AGTCAGTGAATTTCTCATGCTAGAGATCCGATTTATGTTAACCAAGTTT 700
234 ePProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
701 CACCAAGGACCGCGTGTTAACCTGCTTCATTCACATCATGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
751 ATGTTCACTACTCTAGGCTATTTAATTTGTTGTTTCGTTGTAATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
801 AACCAAAATTCGATGAGAAACATTTTAAAACTCTACACAGATTTAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
851 GTACAAGGTATATCTGTACCGACCTTGTTGCAATTCCTCAACAAAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
901 GAATTAATCAATAAATACGATTTGCAAAATTTAGTTGAGATTGCATCTGG 950
317 yGlyIleProLeuSerLysGluIleGlyIleValAlaValAlaArgArgPhe 334
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
951 CGGAGCACCTTTATCAAAAAGAGTTGGTGAAGCTTGTCTAGACGCTTTA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1001 AACTTCCCGGTGTTGCTCAAGGTATAGTTTAAACAGAAAACACATCTGC 1050
351 IleIleIleThrProGluLysAspLysProGlyAlaSerGlyLysVal 367
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1051 ATTATATTATACACAGAAAGAGACGATTAACAGAGAGCTTGTGAAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1101 CGTGCGCTGTTTAAAGCAAAAGTTATTCATCTTGATACCAAAAAATCTT 1150
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1151 TAGCTCCTAACAGACGTCGAGAAAGTTTGTCTTAAAGACCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluGl 417
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1201 AAAGGTTATGTAATATATCCAGAACGCAACAAAGAACTTATTGACGAAGA 1250
417 uGlyTyrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1251 AGGTTGCTGCACACCGGAGATATTTGATATATGATGAGAAAACAATT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1301 TCTTTATGTGCGATCGTTGGAAGTCTTTAATCAAAATACAAAGGATACCA 1350
451 ValProProAlaGluLeuGluLysSerValLeuLeuGlnHisProAsnIleP 467
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1351 GTACCACTGCGCAATTAATGATCCGTTCTTTGCAACATTCATCTATCTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1401 TGATGCTGGTGTTGCCGGGCTTCTGATCCTGATCGAGCTTCAG 1450
484 lValAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1451 GACCCGTTGTTGTAAGGAAAGGAAAGAAATATGTACCGAAAAAGAAAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgLysG 517
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1501 ATGGAATTATGTTCAAGTCAAGTTCAAAATGCAAAAGCTTTAGCTGTGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1551 TGTTGCTGTTGATGAAGTACCTTAAAGGCTTACTGGAAAAATTTGAGC 1600
534 lLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1601 GCAGAGCAATTTAGAAATCTTAAAGAAACAGTTGCTTAAGATG 1644

```


84 erGluAsnCysGluuphepheleProValLeuAlaGlyLeuPheile 100
 251 GTGAAACAGTGAAGAAATCTTATTCCTGATATAGCCGGTTTATTTA 300
 101 GlyValGlyValAlaProThrAsnGluLeuThrLeuArgGluLeuVa 117
 301 GGTGCGGCTGGGCTCCAACTAATGATGATTAACCTACCTGCAATGCT 350
 117 HisSerLeuGlyLeuSerLysProThrIleValPheSerSerLysLysG 134
 351 TCACAGTTTAGGCATCTCTAAGCCACAAATGTTATTTGTTCTAAAAAG 400
 134 LysLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
 401 GATTAGATTAAGTTATATACCTACAAAAAACGGTAACTGCTATTAAC 450
 151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
 451 ATTGTTATATTGGACAGCAAGTGTATATAGAGTTATCATCATCAGCA 500
 167 PAspPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheL 184
 501 CAACCTTATTAATAAAAAACACTCCACAAGGTTCAAGGATCAAGTTT 550
 184 YThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
 551 AAACGTGTAAGATTACCGCAAGAACAGAGTGTCTTATTAATGAACCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLe 217
 601 TCGGGTTCAACCGGTTGCCAAAAAGTGTGCAACTTACATCATGAATTT 650
 217 ValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
 651 GGTACTAGATTCTTCACGCTAGAGATCCAAATTATGAGAAACCAAGTT 700
 234 erProGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
 701 CACAGCAGCAGGCTATTTTAACGTAGTACCATTCATCATGTTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIe 267
 751 ATGTTACTACTTAGGCTATCTAATCTAGTGTGTTTCTGATTCATGTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
 801 AAGGAATTTGACGAAGACACTTTTAAACACACTGCAAGATTACAAAT 850
 284 YSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 851 GTTCAGACGTTATCTTGTACGACCTTGTTCATCTTATATACAACT 900
 301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
 901 GAATTCCTGATTAATATGATTTATCAAAATTAAGTTGAATTCACCTCG 950
 317 yGlyAlaProLeuSerLysGluIleGlyLysAlaValAlaIleArgArgPhe 334
 951 CGAGACACCTTATCTTAAGAAATTCGTACACTGTTGCTAGACGTTT 1000
 334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
 1001 ATTTACCGGGGTCTGCTCAAGGCTATGTTTAAACAGAAACAACCTCTG 1050
 351 IleIleIleThrProGluGlyAspAspLysProGlyValAspSerGlyLysVa 367
 1051 ATTATATTATACACCGAAGGCGATGATTAACAGGAGTCTTCGSCAAAGT 1100
 367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
 1101 TGTGCTATATTAAAGCAAAAGTTATCGATCTGATCTAAAAAACTT 1150
 384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400

1151 TGGCCCGAAGACAGCGGAGAGTTTGTGTAAGGTCCTATGCTTAA 1200
 401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluL 417
 1201 AAAGGTATGTACATATATCCAGAACACACAGAAATATCATGATGAAGA 1250
 417 uGlyTyrPheHisThrGlyAspIleGlyTyrTyrAspGluLysHisP 434
 1251 AGGTGTGTCACACAGAGATGATGTTGATTAACATGAAGAAAAACAT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTyrGlyGln 450
 1301 TCTTTATCGTGATTCCTTGAAGCTTTTAATCAAAATACAAAGATACA 1350
 451 ValProAlaGluLeuGlnSerValLeuGlnHisProAsnIleP 467
 1351 GTACCACTGCTGAATTAATCTGTTCTTTGCAACATCCAAATATTT 1400
 467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLeuP 484
 1401 TGTATCGCGCGTGTGCGGCTCCAGATCCATATACCTGTGAGCTCCG 1450
 484 lValAlaValAlaLeuLysLysGlyLysSerMetThrGluLysGluVal 500
 1451 GAGCTGTGTTGTTACTTAGAAGAGAAATCATATGACTGAAAGAAAGTA 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
 1501 ATGAGATTACGTGTGTCATCAATTTCAAAATGCAAAACCTTTGCGTGG 1550
 517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
 1551 TGTCCGTTTGTGAGCAAGATACCTAAAGTCTCATCTGTAATTTGACG 1600
 534 lLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
 1601 GTAAAGCAATTAGAGAAATGAGAAACCAAGTTGTAAGATG 1644

seq_name: gb_pat:E36480
 seq_documentation_block: DNA PAT 07-FEB-2001
 LOCUS E36480 1644 bp
 DEFINITION Luciferase and method for assaying intracellular ATP by using the same.
 ACCESSION E36480
 VERSION JP 199239493-A/4
 KEYWORDS Luciola lateralis.
 SOURCE Luciola lateralis.
 ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pletergota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Elateriforma; Cantharoidae; Lampyridae; Luciola.
 REFERENCE Noriaki, H.S. M.M. Luciferase and method for assaying intracellular ATP by using the Patent: JP 199239493-A 4 07-SEP-1999;
 AUTHORS KIKROMAN CORP
 JOURNAL OS Luciola lateralis
 COMMENT PN JP 199239493-A/4
 PD 07-SEP-1999
 PF 21-DEC-1998 JP 1998363108
 PR NORIAKI HATTORI, SEIJI MURAKAMI
 PT C12N15/09, C12N9/02, C12Q1/66, C12N9/02, C12R1:19, C12N15/00 CC
 PC /db_xref="taxon:7052"

FEATURES
 source location/Qualifiers
 FH Key (1)..(1644).
 FT CDS location/Qualifiers

BASE COUNT 529 a 262 c 348 g 505 t


```

seq_documentation_block:
LOCUS   AR098258             1644 bp    DNA
DEFINITION   Sequence 13 from patent US 6074859.
ACCESSION   AR098258
VERSION     AR098258.1   GI:12807515
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 1644)
AUTHORS     Hirokawa,K., Kajiyama,N. and Murakami,S.
TITLE       Mutant-type bioluminescent protein, and process for producing the
JOURNAL     Patent: US 6074859-A 13 JUN-2000;
FEATURES
     source             location/Qualifiers
BASE COUNT   528 a      260 c      350 g      506 t
ORIGIN
alignment_scores:
     Quality: 2817.00      Length: 548
     Ratio: 5.150          Gaps: 0
     Percent Similarity: 99.818      Percent Identity: 99.818
alignment_block:
US-09-581-241-4 x AR098258
Align seq 1/1 to: AR098258 from: 1 to: 1644
1 MercGluAsnMetGluAsnAspGluAsnIleValIleTyrGlyProGluProPh 17
1 ATGGAAAACATGGAGAACGATGAAAATATGTGTATGTCTGCTGCAACCAT 50
17 eTyrProIleGluGluSerIleGluAlaIleLeuAlaLeuArgLysMet 34
51 TTACCCATATGAGAGAGGATCTGCTGAGACCAATTCGCAAGATATATGG 100
34 sPaArgTyrAlaLysLeuGluAlaIleAlaIlePheThrAsnAlaLeuThy 50
101 ATCGATATGCAAAACTGGAGCAATTCCTTTACTACGCACTTACCGGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGly 67
151 GTGATTTATACGACCGCAATCTAGAAAATCATGCTGTAGAGA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
201 GCGTTAAAGAAATTATGGTTGGTTGTGATGGAGAAATTCGCTATGCA 250
84 eTyrGluAsnGluGluPhePheIleProValIleuAlaGlyLeuPheIle 100
251 GTGAAACCTGTGAAGATCTTATCTCTGATATAGCGGTTATTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgLysLeu 117
301 GGTGTCGGTGGCTCCACTATGAGATTTACACTCTACGTAATTTGGT 350
117 HisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
351 TCACAGTTTAGCATCTCAAGCAACAATTTGATTAGTTCAAAAAAG 400
134 IleuAspLysValIleThrValGlnLysThrValAlaIleLysThr 150
401 GATTGATTAAGATTATACCTGTAACAAAAACGCTAACCTGTTATAAAC 450
151 IleValIleLeuAspSerLysValAspTyrArgLysTyrGlnSerMet 167
451 ATTTGTTATTTGGACAGCAAGTGGATTTATAGAGTTATCAATCCATGGA 500
167 pAsnPheIleLysAsnThrProGlnIlePheLysGlySerSerPhe 184

```

```

|||||
501 CAACCTTATTAATAAAAAACATCCACAAAGTTTCAAGATCAAGTTTAA 550
184 ySThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAACGTGTAAGATTAAACCGCAAGAAACAAAGTTGCTTATTAATGAAC 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
601 TCGGGTTCAACCGGTTTGGCAAAAGGTGCAACTTACTCATGAAATTT 650
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlnAsnGlnVal 234
651 GGTGATCAAGATTTTCTCAACGCTAGAGATCAATTTATGAAACCAAGT 700
234 eProGlyThrAlaIleLeuThrValValProPheHisGlyPheGly 250
701 CACCAAGCAGCGCTATTTTAACGTGATACCATTCATCATGCTTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMet 267
751 ATGTTACTACTTTAGCTATCTAATCTGATGTTTCTGATTTGTCATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLys 284
801 AACGAATTTGACGAGAGACTTTTAAAAACACTGCAAGATTCANAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAGCGTTATCTTGTACCGACTTGTGCAATTCCTTAATGAAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSer 317
901 GAATTACTCGATAATATGATTATCAATTTACTTGAAATTTGATCGG 950
317 yGlyAlaProLeuSerLysGluIleGlyAlaValAlaArgArgPhe 334
951 CGGAGCACTTATCTTAAGAAATTTGTAAGCTGTTGCTAGACCTTTTA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSer 350
1001 ATTTACCGGGGTTGCTCAAGGCTATGTTTAAACGAAACCACTGCA 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyLysSerGlyLys 367
1051 ATTATATCAACCGGAAGCGATGATAAACAGGTCCTCTGCAAACT 1100
367 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThr 384
1101 TGTGCCATTTATTAAGCAAAAGTTATCGATCTGTGATCAAAAAACT 1150
384 euGlyProAsnArgArgGlyGlyValCysValLysGlyProMetLeuMet 400
1151 TGGGCCGAGACAGACGCTGGAGAACTTTGTAAAGGCTCTTAAGCTTAT 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
1201 AAAGTTATATAGATTAATCCAGAACCAACAAAGAAATCATAGATGAGA 1250
417 uGlyTyrLeuHisThrGlyAspIleGlyTyrGlyAspArgGluLysHis 434
1251 AGGTGGTTGTCACAGAGAGATTTGGGTATTAAGTGAAGAAAAACATT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrGlyGlyTyr 450
1301 TCTTTATCGAGATCTTTGAAGTCTTTAATCAAAATCAAGATATTCAA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIle 467
1351 GTAACACCTGCTGAATTAAGATCTGTTCTTGGACATCCAAATATTTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLysLeuPro 484
|||||

```


[illegible]

167 pasnheilethrlsAsnThrProGlnGlyPheLysGlySerPheL 184
 |||||
 501 CAACCTTTATTAACCAACACCTCCACAAAGTTTCAAGATCAAGTTTAA 550
 184 ysthrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
 |||||
 551 AAACGTAGAAAGTTAACCGCAAGCAAGATTGCTTATATGAACT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGlnAsnLe 217
 |||||
 601 TCGGGTTCAACCGGTTTCCAAAGAGTGTGCACACTTACTCATGAAGATGC 650
 217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
 |||||
 651 AGTCACTATATTTTCCACAGTAGAGATTCATTTATGGAACCAAGTTT 700
 234 ePrroGlyThrAlaIleLeuThrValValProPheHisHisGlyPheGly 250
 |||||
 701 CACACAGCGACGGCTATTTTAACTGATGATCAATCCATCATGTTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
 |||||
 751 ATGTTTACTACTTATAGCTATCTACTACTGTGTTTCCATTTGTCATGTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
 |||||
 801 AACGAATTTGACGAAGAGACTTTTAAAAACACACGACAGATTACAAAT 850
 284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
 |||||
 851 GTTCACAGCGTATTTCTGTACCGACTTGTGTTCCATTTCTTAATAGAACT 900
 301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
 |||||
 901 GAATTAAGTAAATATGATTTATCAAAATTGATTGAAATTCATCTGG 950
 317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
 |||||
 951 CGGAGCACCTTATCTTAAGAAATGCTGAGAGCTGTTGCTAGACGTTT 1000
 334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGlnThrSerAla 350
 |||||
 1001 ATTATCCGGGTTCGTCAGAGCTATGCTTAAACAGAAACACCTCTGCA 1050
 351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
 |||||
 1051 ATTATATATCACACCGAGCGAGATGATTAACACAGCTCTTGGCAAAAGT 1100
 367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
 |||||
 1101 TGTCCCATTTATTTAAACCAAAAGTTATCGATCTTGATACAAAAAACTT 1150
 384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
 |||||
 1151 TGGGCCCGAAGACAGCTGAGAAAGTTGTGTAAAGCGTCTCTATGCTTATG 1200
 401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluG 417
 |||||
 1201 AAAGTTATGATGATTAATCCAGAAAGCAACAGAAATTCATAGTGAAGA 1250
 417 uGlyTyrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
 |||||
 1251 AGGTTGGTGCACACAGAGATATTTGGTATTTAGCAGCAAGAAAAACATT 1300
 434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
 |||||
 1301 TCTTATATCGATCGATTGAAAGCTTTTAATCAATAACAAAGGATATCAA 1350
 451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
 |||||
 1351 GTACCACTGCTGGAATTTGAATCTGTCTTTTGCACACATCAAAATATTTT 1400

467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
 |||||
 1401 TGATGCCGGCGTTCGCTGCGTCCAGATCTTATAGCTGGGAGCTCCGG 1450
 484 lValAlaValAlaLeuLysLysGlyLysSerMetThrGlyLysGluVal 500
 |||||
 1451 GAGCTGTTGTTGTTACTTGAAAAAGGAAATCTATGACTGAAAAAGATTA 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG 517
 |||||
 1501 ATGATTACGTTGCTAGTCAAGTTTCAATGCAAAAGCTTTGGTGATGG 1550
 517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
 |||||
 1551 TGTCCGTTTGTGACGACAGACTTAAGGCTCTACGTTGTTAAATTTGACG 1600
 534 lLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
 |||||
 1601 GTAAAGCAATTAGAGAAATACTGAAGAAACAGTTGCTTAAGATG 1644
 seq_name: gb_in:LLDUCI

seq_documentation_block: 1781 bp mRNA
 LOCUS LLDUCI
 DEFINITION L. lateralis luciferase mRNA.
 ACCESSION X66919.38389
 VERSION X66919.1 GI:9526
 KEYWORDS luciferase
 SOURCE Japanese firefly.
 ORGANISM L. lateralis
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 Elateriformia; Cantharoidae; Lampyridae; Luciola.

REFERENCE
 AUTHORS Tatsumi, H., Kajiyama, N. and Nakano, E.
 TITLE Molecular cloning and expression in *Escherichia coli* of a cDNA
 MEDLINE clone encoding luciferase of a firefly, *Luciola lateralis*
 92305054
 JOURNAL Biochim. Biophys. Acta 1131 (2), 161-165 (1992)
 REFERENCE 2 (bases 1 to 1781)
 AUTHORS Tatsumi, H.
 TITLE Direct Submission
 JOURNAL Submitted (21-JUL-1992) H. Tatsumi, Research and Development
 Division, Kikkoman Corporation, 399 Noda, Noda City, Chiba 278,
 JAPAN

FEATURES
 1. 1781
 location/Qualifiers
 /organism="Luciola lateralis"
 /db_xref="taxon:7052"
 59..1705
 /codon_start=1
 /product="luciferase"
 /protein_id="CAA47358.1"
 /db_xref="GI:9527"
 /db_xref="SWISS-PROT:O01158"
 /translation="MEMMENDENTVVGPEPEFYPIEGSAGAO LRKYMDRYAKGAIAF
 GNAITGVDYTYAETLEKSCCLGEALIKYGVVGRALASGNEEEFIPYLAGLFGV
 TVALPTNEYTLRELVHSLSKPTIVSSKKGLDKYTVQKTYAIKTIYILDSKDY
 RGYSDMNFIRKNTPOGKSSSEFVAVNRKEOVALIMNSGSTGLPKGQULHENAV
 TRFSHADPIYGNQVSPGTALIVAPPHGEGMTTGLVLTGFRIVMLTKFDEETFL
 KTLDDYKSSVILPTLFAILINRSELDKDYLSNLVJASGAPLSKEIEAVARRFN
 LRGVROGGLRETSALIIIPPEGDKRGASGVVPLFKAKVIDLTSTKTEGPNRRGV
 CVKRPMLMKGVNDPEATRETIDEEGMLHNGDIGYDEEKHFFIVDLKSLIKYKGV
 VPPAELEVLQHNINFDAGYAGVPPDIAPETLPGAVVLEKGSMTKEKVMDDVVASGV
 SNARKLRGVFVDEVPKGLKIDGKIRIRILKPKVAKK"

CDS

BASE COUNT 579 a 276 c 367 g 559 t
 ORIGIN

alignment_scores: Length: 548
 Quality: 2814.00 Gaps: 0
 Ratio: 5.144
 Percent Similarity: 99.818 Percent Identity: 99.635

alignment_block:
 US-09-581-241-4 x LLLUCI ..
 Align seg 1/1 to: LLLUCI from: 1 to: 1781

```

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
  |||||||
59 ATGGAAACATGGAGACGATGAAAATATGTGTATGTCTCTGAACCAT 108
  |||||||
17 eTyrProIleGluGluGlySerAlaGlyAlaGluLeuArgLysTyrMet 34
  |||||||
109 TTACCTTATGAGAGGATCTGCTGGAGCAATGCGCAAGTATATGG 158
  |||||||
34 sPaqTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
  |||||||
159 ATCGATATGCAAAACTTGGAGCAATGCTTTACTAACGGCACTTACCGGT 208
  |||||||
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyG 67
  |||||||
209 GTCGATATACGTACGCGGAATCTTAGAAAAATCATGCTGTATAGGAGA 258
  |||||||
67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuGly 84
  |||||||
259 GCGTTTAAAGATATATGTTGGTTGGTTGATGAGAAATTCGCTTATGCA 308
  |||||||
84 eArgLysAsnGlyGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
  |||||||
309 GTGAAACTGTGAAGAAATCTTTATCTGTATATAGCCGGTTATATATA 358
  |||||||
101 GlyAlaGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
  |||||||
359 GGTTGCGGTGGCTCCAACTAATGAGATTACACTCTACGTGAATGGT 408
  |||||||
117 LHisSerLeuGlyLysSerLysProThrIleValLpHeSerSerLysLys 134
  |||||||
409 TCACAGTTTAGGCATCTTAAGCAACAATGTATTAGTTCTAAAAAAAG 458
  |||||||
134 LysLeuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150
  |||||||
459 GATTAGATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGT 508
  |||||||
151 IleValIleLeuAspSerLysValAspTyrArgLysGlyThrGlnSerMet 167
  |||||||
509 ATGTATATATTTGAGACAGCAAAAGTGAATTAAGGTTATCAATCCATGGA 558
  |||||||
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
  |||||||
559 CAACCTTATTAAAAAACAACACCTCCACAGGTTTCAAGATCAAGATTGTA 608
  |||||||
184 yThrValGluValAsnArgLysGluGluValAlaLeuIleMetAsnSer 200
  |||||||
609 AACTGTAGAGTTTAAACCCCAAGAACAAAGTGTCTTTATTAATGAACCTCT 658
  |||||||
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLe 217
  |||||||
659 TCGGGTTCAACCGGTTGGCCAAAAGGTGTCAACTTACTCATGAATAATGC 708
  |||||||
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGluValS 234
  |||||||
709 AGTCACTAATATTTTCTCAGCGTAGAGATCAATTAATGAAACCAAGTTT 758
  |||||||
234 eArgProGlyThrAlaIleLeuThrValAlaProPheHisHisGlyPheGly 250
  |||||||
759 CACAGGCGACGGCTATTTAACTAGTACATTCATTCATCATGTTTGGT 808
  |||||||
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
  |||||||
809 ATGTTTACTACTTTAGGCTATCTAATCTGTGTGTTTCGATATGTCATGTT 858
  |||||||
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
  |||||||
859 AACGAAATTTGACGAGAGACTTTTAAAAACACTGCAAGATTACAAAT 908
  |||||||

```

```

284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
  |||||||
909 GTTCACACCTTATTTCTGTACCGACTTGTTCATCTTATATTAACACT 958
  |||||||
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
  |||||||
959 GAATTAATCATTAATATGATTTATCAAAATTAATGAAATTCACATCTCG 1008
  |||||||
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
  |||||||
1009 CCGAGACACCTTATTTAAAGAAATTCGTACAGCTGTGTCTACACGTTTA 1058
  |||||||
334 sLeuProGlyValAlaArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
  |||||||
1059 ATTTACCGGGGTTCGTCAAGGCTATGCTTAAACGAAACCAACTCTGCA 1108
  |||||||
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
  |||||||
1109 ATTATTAATCACACCGAAGCCATGATTAACCAAGGTGCTTGGCAAGT 1158
  |||||||
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
  |||||||
1159 TGTGCATATTTTAAAGCAAAAGTTATCGATCTGTGATCTAATAAAACTT 1208
  |||||||
384 eGlyProAsnArgArgGlyGlyValLysValLysGlyProMetLeuMet 400
  |||||||
1209 TGGCCCCCAACAGACGTGAGAAATTTGTGTAAGGGTCTATAGCTATG 1258
  |||||||
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluG 417
  |||||||
1259 AAAGTTATATGTAGATTAATCCAGAACCAACAGAAATCATATGATGAAG 1308
  |||||||
417 uGlyThrPheHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
  |||||||
1309 AGGTGTGGTGCACACAGAGATTTGGTATTACGATCAAAAAAACACTT 1358
  |||||||
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
  |||||||
1359 TCTTATATCGTGAATGTTGAAGTCTTTAATCAATACAAAGATATCA 1408
  |||||||
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
  |||||||
1409 GTACCACTGCTGAATTAAGTATGTTCTTTGCAACATCCAAATATATTT 1458
  |||||||
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
  |||||||
1459 TGATCCCGCGGTGCTGTGCGCTCCAGATCTATAGCTGGTGAGCTTCGG 1508
  |||||||
484 lValAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
  |||||||
1509 GAGCTGTGTGTACTTGAAGAAAGAAATCTATGACTGAAAAAAGAGTA 1558
  |||||||
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG 517
  |||||||
1559 ATGATTAATACGTGCTAGTCAAGTTCATTAATGCAAAAGTGTGCTGTGG 1608
  |||||||
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
  |||||||
1609 TGTCGCTTTGTGAGCAGAGTACTTAAGGTCTACACGTGAAAAATTCAGC 1658
  |||||||
534 lLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
  |||||||
1659 GTAAAGCAATTTAGAGAAATCTGAGAAACCAAGTGTCTAGATG 1702
  |||||||

```

seq_name: gb_pat:AR043317

seq_documentation_block:

LOCUS AR043317 1908 bp DNA

DEFINITION Sequence 8 from patent US 5814465.

ACCESSION AR043317

VERSION AR043317.1 GI:5964325

KEYWORDS

PAT 29-SEP-1999

501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
1501 ATGGATTACGTTGCTAGTCAAGTTCAATGCAAAACGTTTCGGTG 1550
517 ValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
1551 TGTCCGTTTGTGGACGAGTACCTTAAGGCTCAGCTGTAATAATTGACG 1600
534 LysValAlaIleArgGluIleLeuLysLysProValAlaLys 547
1601 GTAAAGCAATTAGACAAATGAGAAACCGTGTAG 1641
seq_name: gb_pat:AR062709

seq_documentation_block:
LOCUS AR062709 1908 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 8 from patent US 5843746.
ACCESSION AR062709
VERSION AR062709.1 GI:5990400
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1908)
AUTHORS Tatsumi, H., Fukuda, S., Kikuchi, M. and Koyama, Y.
TITLE Biotininated firefly luciferase, a gene for biotininated firefly luciferase, a recombinant DNA, a process for producing biotininated luciferase and a bioiminescent analysis method
JOURNAL Patent: US 5843746-A 8 01-DEC-1998;
FEATURES
source 1. 1908
BASE COUNT 598 a 329 c 426 g 555 t
ORIGIN

alignment_scores:
Quality: 2814.00 Length: 547
Ratio: 5.144 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.817

alignment_block:
US-09-581-241-4 x AR062709 ..

Align seg 1/1 to: AR062709 from: 1 to: 1908

1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProph 17
1 ATGGAAGAAACATGTGAGAACGATGAAATATGTGTATGCTGCAACCAT 50
17 eTyrProIleGluGlySerAlaGlyAlaIleAlaPheThrAsnAlaLeuThrGly 34
51 TTTCCCTATTGAGAGGAGATCTCTGAGACACAATTGGCGCAAGATATG 100
34 sPArGTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 ATCGAATATGAAAACTTGAGCAATTGCTTTACTAACGCACCTTACCGGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLysLeuGly 67
151 GTGCAATTATACGTCGCGCAATCTTAGAAAAATCATGCTGTAGAGAGA 200
67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuCys 84
201 GGCTTTAAAGAAATTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 250
84 eGluAsnCysGluGluPheIleProValAlaLeuAlaGlyLeuPheIle 100
251 GTGAAAGAACTGTGAAGAGTCTTATCTGTATTAACCGGTTATTTATA 300
101 GlyValAlaGluValAlaProThrAsnGluIleTyrThrLeuArgGluLeu 117
301 GGTGTGCTGTGCTCCAACTAATAGATTATACACTACGTGGAATTTGGT 350

117 HisSerLeuGlyIleSerLysProThrIleValPheSerLysLysG 134
351 TCACAGTTTGGCATCTCTAACCAACAATGCTATTTAGTTCTTAAAAAG 400
134 LysLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
401 GATTGATTAAGTTATTAAGTACAAAAACGTAACGCTATTTAAAAAC 450
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
451 ATGTATATATGGACAGCAAGTGATATAGAGTTATCAATCATGCA 500
167 PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerPhe 184
501 CAACCTTATTTAAAAAAACCTCCCAAGGTTTCAAGATCAAGTTTAA 550
184 ySthrValAlaValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAACGTGAAAGTTAACCGCAAGACAAAGTGGCTTATATATGAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
601 TCGGGTTCACCGGTTTGGCCAAAAGGTGTCACACTTACTCATGAATTT 650
217 uValThrArgPheSerHisAlaArgAspProIleTyrGlyLysGlnVal 234
651 GGTACAGCGTTTCTCAGCGCTAGAGATCCAAATTATGGAACCAAGTT 700
234 eProGlyThrAlaIleLeuThrValAlaProPheHisIleGlyPheGly 250
701 CACGAGCAGCGTATTTTAACTGATGACCATTCATCATGCTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMet 267
751 ATGTTACTACTTTAGGCTATCACTGCTGCTTTGCTATGTCATGCT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGluAspTyrLysC 284
801 AACGAATTTGACGACAGACACTTTTAAAAACCTGCACATTCATAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAACGCTTATCTGTACCGACTTGTGCAATTCCTTAATAGAA 900
301 GluLeuLeuAspLysTyrAspLysSerAsnLeuValGlnIleAlaSerG 317
901 GAATTAATCATTAATATGATTTATCAAAATTTAGTTGAAATGCAATCT 950
317 yGlyAlaProLeuSerLysGluIleGlyAlaValAlaArgPhe 334
951 CGGAGCACCTTATCTTAAAGAAATGGTGAAGCTGTTGTAAGCTTTA 1000
334 sLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
1001 ATTACCGGGTGTGCTCAAGGCTATGTTAAACAGAAACCTCTGCA 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLys 367
1051 ATTATATATCACACCGAAGCGATGATTAACAGAGTCTTCTGCAAG 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThr 384
1101 TGTGCATTTATTTAAAGCAAAAGTTATCGATCTTGTACTAAAAAACT 1150
384 eGluLysProAsnArgArgGlyValCysValLysLysProMetLeuMet 400
1151 TGGGCCCGAAGACAGCTGGAAGATTGTGTAAAGGCTCTTATGCTTA 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleLeuAspGlu 417
1201 AAAGTTATGTAGATTAATCCAGAGCAACAAGAAATCATGATGATA 1250

417 uclytrpleuisthrghlyaspilleglytyrtyraspgluulyshtsp 434
 1251 agcttggtgacacagacagatattggctattgacagcaaaaaacatt 1300
 434 hephellvalaspargleuyserserleuilelytyrlysglytyrigin 450
 1301 tctttatcgatcgatcgctttgcaagcttttaactcaaatcacaaagcatatcaa 1350
 451 valproalagluleugluservalleuenglnhisproanlleph 467
 1351 gtagccacctgctgaattagaaatctcttttgcaacatccaaaataatttt 1400
 467 easpalaglvalalaglvalproasproilleagllyluenuprog 484
 1401 tgatgccggcgttgctggcgttcacatctatgactgtgtaacgttccgg 1450
 484 lyalvalvalvalleuylslysglytyrsermetthrlylulysgluval 500
 1451 gacgtgcttcttacttgcaaaaaagcaaatctgactgcaaaaaagacta 1500
 501 metaspyrvalalaserclnvalserasnalyasargleuarglygl 517
 1501 atggattacgttgctagcaaatcttcaaatgcaaaacgttgcgctgg 1550
 517 yvalargphevalaspgluvalprolysglyleuthrglyllysileaspg 534
 1551 tggcgttgggttgagagagatgactgaaggttcaactgtgtaaaattgacg 1600
 534 lylysallaleargluileuylslysglytyrprovalalalys 547
 1601 gtaaaacatttgagaaatgactgaagaaacacagttgctmag 1641

seq_name: gb_pat:E12279

seq_documentation_block:

LOCUS E12279 1908 bp DNA PAT 24-JUN-1998
 DEFINITION cDNA encoding biotinylated luciola luciferase.
 ACCESSION E12279
 VERSION E12279.1 GI:3251113
 KEYWORDS JP 1996308578-A/4.
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 Elateriformia; Cantharidea; Lampyridae; Luciola.
 Luciola lateralis.
 Luciola lateralis

REFERENCE
 AUTHORS
 TITLE
 1 (bases 1 to 1908)
 Tatsumi,H., Fukuda,M., Kikuchi,M. and Koyama,T.
 BIOTIN FIRE FLY LUCIFERASE, BIOTIN FIRE FLY LUCIFERASE GENE, NEW
 RECOMBINANT DNA, PRODUCTION OF BIOTIN FIRE FLY LUCIFERASE AND
 BIOLUMINESCENCE ANALYSIS
 Patent: JP 1996308578-A 4 26-NOV-1996;

JOURNAL
 KIKKOMAN CORP
 OS
 Luciola lateralis
 PN JP 1996308578-A/4

COMMENT

PD 26-NOV-1996
 PF 24-APR-1995 JP 1995098857
 PR 27-JUL-1994 JP 94P 193798, 14-MAR-1995 JP 95P 54625 PI
 TATSUMI HITOKI, FUKUDA MASARU, KIKUCHI MAMORU, KOYAMA TAJIJI PC
 C12N15/09,C12N9/02,C12Q1/26,(C12N9/02,C12R1:185); CC
 strandedness: Double;
 CC topology: Unknown;
 CC hypothetical: No;
 CC anti-sense: No;

key location/Qualifiers

FT source 1. 1908 /organism='Luciola lateralis' FT

FT /clone='pHLf248', 1. 1908
 FT mat-peptide /product='biotinylated Luciola luciferase'.
 FT Location/Qualifiers
 1. 1908
 /organism='Luciola lateralis'

BASE COUNT 598 a 329 c 426 g 555 t
 ORIGIN
 alignment_scores:
 Quality: 2814.00 Length: 547
 Ratio: 5.144 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.817
 alignment_block:
 US-09-581-241-4 x E12279 ..
 Align seq 1/1 to: E12279 from: 1 to: 1908

1 MetcLusnmetgluasnaspgluasnllevaltyrlyglyprogluproph 17
 1 ATGGAAAAACATGAGAGAGATGAAAAATATGTGTATGCTCTGAACCAT 50
 17 etyrproilleglugluclyseralaglvalagluleuarglytyrmeta 34
 51 TTACCCATTGTAAGAGGATCTGCTGGACACAATTGCGCAAGTATATG 100
 34 sparqyrvalalysleuclvallealaphethrasnalaleuthrgly 50
 101 ATCGATATGCAAAACCTTGAGCAATTCCTTTACTTAACGACACTACCGGT 150
 51 Valaspyrthrtyrvalaglulthrleuglulyssercysleuglygl 67
 151 GTCGATTATACGTACGCCGAACTACTTGAATAATCATGCTGTACAGAGA 200
 67 uAlaleuylsasntyrlyleuvalalaspglyargllealeucys 84
 201 GCGTTTAAAGAAATATGTTGTTGTTGATGAGAAATGCGTTATGCA 250
 84 ergluasrcysglugluphehelleprovalleualaglyleuphele 100
 251 GTGAAMACTGGAAGAGTCTTATTCCTGTATTAACCGGTTATTTATA 300
 101 GlyvaliglyvalalaprothrAsngluiletyrthrleuarglualeu 117
 301 GGTGCGGTGGCTGCCAATCATGATGATTAACCTGACTGAATGGT 350
 117 IhisSerleuglylserlysprothrIlevalpheserlyslysg 134
 351 TCACAGATTAGGACATCTTAAGCCCAATGTATTTAGTTCTAAAAAG 400
 134 lyLeuAspLysValIlethrValglnlysthrvalIlethr 150
 401 GATTGATTAAGTTATTAACGTACAAAACGGTACTGCTATTAAAAAC 450
 151 lleValIleLeuAspSerLysValAspTyraRgLyTyrgInSerMetas 167
 451 ATTGTATATTGGACAGCAAGTGATTAAGAGTTATCAATCATGGA 500
 167 pasnhelelyllysasnthrProglnglypheylslyserSerphel 184
 501 CAACCTTATTAATAAAAAACACCCCAAGTTTCAAGGATCAAGATTTPA 550
 184 ystThrValgIuValAsnArgLysGluInValAlaLeuIleMetasSer 200
 551 AAACGTAGAACTTAACCCCAAGCAAGATGCTCTTATATAGAACCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuthrHisGluAsnle 217
 601 TCGGGTTCAACCGGTTTGCACAAAGGTGCAACTTACTCATGAATAATT 650
 217 uvalthrargpheserHisAlaArgasprollethyrglyAsngInValS 234
 651 GGTACAGCGCTTTTCTCACAGCTAGAGATCCAAATTATGGAACCAAGTTT 700
 234 erProGlyThrAlaIleLeuthrValalProphHisHisGlyphegly 250

```

701 CACGAGCAGCGCTATTTAACTGATACCATTCATCATGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIle 267
|||||
751 ATGTTTACTTACTTAGGCTATCTACTGTTGTTTGGTATGTCATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
|||||
801 AACCAAAATTTGACGAGAGACTTTTAAAAAACAACCTGCAAGATTACAAAT 850
284 ySerSerValIleLeuValProPheThrPheAlaIleLeuAsnArgSer 300
|||||
851 GTTAACGCTTATCTGTACCGACTTGTGTGCAATTTCTTAATAGAAAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleLeuAsnSer 317
|||||
901 GAATTACTTCGATTAATGATGATTATCAAAATTGATTGCAAAATTCATCTGG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheA 334
|||||
951 CGGAGCACCTTATCTAAAGAAATGGTGAAGCTGTGCTAGACGTTTGA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrPheSerAla 350
|||||
1001 ATTTACCGGGTGTTCGTCAGAGCTATGTTTAACAGAAACAACCTCTGCA 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVal 367
|||||
1051 ATTATTAATACACCGGAGGCGATGATTAACACGAGTCTTGCAAAAGT 1100
367 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
|||||
1101 TGTGCCATTTTAAAGCAAAAGTTATCGATCTGTGATCTAAAAAAAAGCTT 1150
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
|||||
1151 TGGGCCCCGAAACAGACGCTGGAGAGTTGTGTAAAGGCTCTATGCTTAG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
|||||
1201 AAAGGTATGATGATATCCAGAACACACAGAAATCATGATGAAACA 1250
417 uGlyTyrPheHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
|||||
1251 AGGTGTGTTGCACACAGAGATATTTGGTATTAACGATGAGAAAAACATT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
|||||
1301 TCTTATTCGTGATCGTTGAAGCTTTAATCAATAACAAAGATATACAA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePhe 467
|||||
1351 GTACCCCTGCTGCAATTAAGAAATCTGTTTGGCAACATCCCAAAATATTTT 1400
467 eaSpAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
|||||
1401 TCGATGCCGCGTGTGCGCTCCAGATCTTACTGCTGAGCTTCCGG 1450
484 ValAlaValAlaValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
|||||
1451 GAGCTGTGTGTACTTGAAGAAAGAAATCTATGACTGAAAGAAAGAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
|||||
1501 ATGGAATTACGTTGCTAGTCAAGTTCCAATGCAAAACGTTTGGCGGTGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysLysLeaspG 534
|||||
1551 TGTCCGTTTGTGGAGAGTAAAGTCTCACTGCTAAATTTGACG 1600
534 yLysAlaIleArgGluIleLeuLysLysProValAlaLys 547
|||||
1601 GTAAAGCAATTAGAGAAATACTGAGAAACACAGTGTAG 1641

```

```

seq_name: gb_pat:E16288
seq_documentation_block:
LOCUS E16288 1920 bp DNA PAT 28-JUL-1999
DEFINITION DNA encoding fusion protein which comprises luciferase and
lysostaphin.
ACCESSION E16288
VERSION E16288.1 GI:5710971
KEYWORDS JP 1998150991-A/1.
SOURCE
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 1920)
AUTHORS Tatsumi,H., Fukuda,M. and Nagahara,A.
TITLE LUCIFERASE-LYOSTAPHIN FUSED PROTEIN, ITS PRODUCTION AND
JOURNAL BIOLUMINESCENT ANALYSIS USING THE SAME
PATENT: JP 1998150991-A 1 09-JUN-1998;
KIKKAWAN CORP
COMMENT
OS None
OC Artificial sequences.
PN JP 1998150991-A/1
PD 09-JUN-1998
PF 25-NOV-1996 JP 1996328042
PI TATSUMI HIROKI, FUKUDA MASARU, NAGAHARA AYUMI PC
C12N15/09,C07H21/04,C07K19/00,C12N9/02,C12N9/52,C12P21/02,PC
C12Q1/66
PC (C12N9/02,C12R1:19),(C12N9/52,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..1920
FT location/Qualifiers
1..1920
/organism="Artificial sequences",
/organism="unidentified"
/db_xref="taxon:332644"
BASE COUNT 627 a 310 c 407 g 576 t
ORIGIN
alignment_scores:
Quality: 2814.00 Length: 547
Ratio: 5.144 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.817
alignment_block:
US-09-581-241-4 x E16288 ..
Align seg 1/1 to: E16288 from: 1 to: 1920
1 MetLuAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProH 17
|||||
1 ATGGAACAAATGAGAGAGATGAAATATATGTGTATGCTGCTGAACCAAT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnIleuArgLysTyrMetA 34
|||||
51 TTACCCATTGGAAGAGGATCTGCTGGAGCGCAATTTGCGCAAGTATATGG 100
51 TTACCCATTGGAAGAGGATCTGCTGGAGCGCAATTTGCGCAAGTATATGG 100
34 spArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
|||||
101 ATCGATTATGCCAAAACCTTGAGCAATTTGCTTTACTTAACGACTTACCGGT 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysCysLeuGly 67
|||||
151 GTCGATTATACGACCGCGCAATCTTAGAAAATATCAATGCTGTCTAGAGA 200
67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuCys 84
|||||
201 GCGTTTAAAGAAATTTATGCTTTGCTTTGATGGAAGAAATTTGCTATGCA 250

```

```

84  erGLuAsnCysGLuGluNhepHeilLeProValLeuAlaGlyLeuPheIle 100
|||||
251  GTGAAGAACTGTGAAGAGTTCTTTATCTGATTAAGCCGGTTATTTATA 300
|||||
101  GLyValIGlyValAlaProThrAsnGluIleThrThrLeuArgGluLeuVal 117
|||||
301  GGTGCGGCTGTGGCTCAACTAATGAGATTACACTGACTGAGTGAATGGT 350
|||||
117  LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysGly 134
|||||
351  TCACAGTTTAGGCATCTCTAGCCACAAATGTATTAGTTCTAAAAAG 400
|||||
134  ILeuAspLysValIleThrValGlnLysThrValThAlaIleLysThr 150
|||||
401  GATTAGATTAAAGTTAATCTGACAAAAACGGTAACTGCTATTAAAAAC 450
|||||
151  ILeuValIleLeuAspSerLysValAspLysArgLysGlyArgLysSerMetAs 167
|||||
451  ATTGTTATATTGGACACGAAAGTGAATTAAGAGTTATCAATCCATGGA 500
|||||
167  PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
|||||
501  CAACCTTTATTTAAAAAACACATCCCAAGGTTTCAAGGATCAAGTTTAA 550
|||||
184  ysthValIGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
|||||
551  AAACGTGTAGAGTTAACCGCAAGAACAAAGTGTCTTTATTAATGAACTCT 600
|||||
201  SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLe 217
|||||
601  TCGGGTTCAACCGGTTTGGCAAAAGGTGCAACTTACTCATGAATAATT 650
|||||
217  uValThrArgPheSerHisAlaArgAspProIleThrGlyAsnGlnValS 234
|||||
651  GGTCAAGCGCTTTTTCACGCTAGAGATCAATTAATGAAACCAAGTTT 700
|||||
234  expProGlyThrAlaIleLeuThrValAlaProPheHisHisGlyPheGly 250
|||||
701  CACGAGGACGCGTATTTTAACTGATGATCCATTCATCATGTTTGGT 750
|||||
251  MetPheThrThrLeuGlyThrLeuThrCysGlyPheArgIleValMetLe 267
|||||
751  ATGTTTACTACTTATGAGTATCTAACTGTGGTTTCTGATTTGTCATGTT 800
|||||
267  uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspLysLysC 284
|||||
801  AACGAAATTTGACGAAGAGACTTTTAAAAACACTGCAAGATTACAAAT 850
|||||
284  ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
|||||
851  GTTCAAGCGTATTCTGTATACGACCTTTGTTTCAATTTCTTATAGAACG 900
|||||
301  GluLeuLeuAspLysLysAspLeuSerAsnLeuValGluIleAlaSerG 317
|||||
901  GAATTTACTCGATAAATATGATTTATCAAAATTGTTGAATTTGCACTCG 950
|||||
317  yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
|||||
951  CGGAGCACCTTTTATCTTAAGAAATTTGTTGAGAGTGTTCAGACGTTT 1000
|||||
334  snLeuProGlyValArgGlnGlyThrGlyLeuThrGluThrThrSerAla 350
|||||
1001  ATTTACCGGCTGTTGCTCAAGGCTATGTTTAAACAGAAACAACTCTGCA 1050
|||||
351  IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
|||||
1051  ATTATTTATCACACCGGAGGCGATGATTAACCAAGGTCCTTCTGCAAGT 1100
|||||
367  IValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThr 384
|||||
1101  TGTGCGCATTTATTAAGCAAAAGTTATCGATCTTGATACTAAAAAACTT 1150

```

```

384  euGIyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
|||||
1151  TGGGCCCGAAGACAGACGTCGGAGAACTTTGTCTAAAGGCTCTTATGCTTAG 1200
|||||
401  LysGlyThrValAlaAspAsnProGluAlaThrArgGluIleLeuAspGlu 417
|||||
1201  AAAGTTATGATGATTAATCCAGAAAGCAAGAAAGAAATCATGATGTAAG 1250
|||||
417  uGlyThrLeuHisThrCysLysPheGlyThrLysArgPheGluLysHisP 434
|||||
1251  AGGTTGGTTCACACAGAGATATTTGGTATTTACGATGAAGAAAAACATT 1300
|||||
434  hePheIleValAspArgLeuLysSerLeuIleLysThrLysGlyThrGln 450
|||||
1301  TCTTTATCGTGATCGCTTTGAAGCTTTATTCAAATACAAAGATATTCAC 1350
|||||
451  ValProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePhe 467
|||||
1351  GTACCACCTGCTGAATTAGAAATGTTCTTTTGGCAACATCCAAATATTTT 1400
|||||
467  eaSPAlaGlyValAlaGlyValProAspProIleAlaGlyIleuProG 484
|||||
1401  TGATGCGGCGCTGCTGCGCTTCCAGATCCTATAGCTGCTGAGCTTCGG 1450
|||||
484  IyAlaValAlaValAlaLeuLysLysGlyLysSerMetThrGluLysGluVal 500
|||||
1451  GAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 550
|||||
501  MetAspLysValAlaSerGlnValSerAsnAlaLysArgLeuArgLysG 517
|||||
1501  ATGATTAAGTTAGTTGATGCAAGTTCAAAATGCAAAAGTTGCTGCTGG 1550
|||||
517  yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
|||||
1551  TGTGCGTTTGTGGACGAGTAACTAAAGGCTCAGCTGCTGCTGCTGCTG 1600
|||||
534  IyLysAlaIleArgGluIleLeuLysLysProValAlaLys 547
|||||
1601  GTAAAGCAATTAGAGAAATTAAGTAAAGAAACCAAGTTGCTTAAG 1641
|||||
seq_name: gb_pat:AR043316
seq_documentation_block:
LOCUS AR043316 1704 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 5 from patent US 5814465.
ACCESSION AR043316
VERSION AR043316.1 GI:5964324
KEYWORDS
SOURCE
ORGANISM Unknown.
UNCLASSIFIED.
REFERENCE 1 (bases 1 to 1704)
AUTHORS Tatsumi,H., Fukuda,S., Kikuchi,M. and Koyama,Y.
TITLE Biotinated firefly luciferase, a gene for biotinated firefly
luciferase, a recombinant DNA, a process for producing biotinated
luciferase and a bioluminescent analysis method
JOURNAL Patent: US 5814465-A 5 29-SEP-1998;
FEATURES
source location/Qualifiers
BASE COUNT 541 a 275 c 361 g 527 t
ORIGIN
alignment_scores:
Quality: 2801.00 Length: 546
Ratio: 5.130 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.451
alignment_block:
US-09-581-241-4 x AR043316 ..
Align seg 1/1 to: AR043316 from: 1 to: 1704

```

3 AsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrPr 19
 : : : : :
 67 AGTCGAGACGATGAAATATGTGTATGTCTCTGACCATTTTACC 116
 19 cllcglugllySerAlaGlyAlaGlnLeuAlaGlyTyrMetAspArg 36
 : : : : :
 117 TATTGAGAGGATCTGCTGAGACCAATTCGCAAGATATGTATGATGAT 166
 36 yralAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValAsp 52
 : : : : :
 167 ATGCAAAACTTGAGCAATTCCTTTTACTACACGACTTACCGGTGTGAT 216
 53 TyrThrTyrAlaGluTyrLeuGlyLysSerCysLysLeuGlyAlaLe 69
 : : : : :
 217 TATACTGACCCGCAATCTAGAAAATTCATGCTGTAGAGAGCGCTT 266
 69 uLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCysSerGlu 86
 : : : : :
 267 AAAGATATATGCTTGTGTGTGATGAGAAATTCGCTTATGACGTGAAA 316
 86 sncysglugluPhePheIleProValLeuAlaGlyLeuPheIleGlyVal 102
 : : : : :
 317 ACTGTGAGAGATCTTATTCCTGATATACCGGTTATTTATATAGTGTCT 366
 103 GlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHis 119
 : : : : :
 367 GGTGTGCTCCAACTAATGAGATTTACACTCTACGTGAATTTGGTTCAG 416
 119 rleuGlyIleSerLysProThrIleValPheSerSerLysLysGlyLeu 136
 : : : : :
 417 TTTAGCATCTCTAAGCCACAACTTGTATTTAGTCTAAAAAAGATTAG 466
 136 spLysValIleThrValGlnLysThrValThrAlaIleLysThrIleVal 152
 : : : : :
 467 ATTAAGTTATTAAGTGCACAAAACGCTTAACTCTATTTAAACCATGTT 516
 153 IleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAspAsn 169
 : : : : :
 517 ATATTGACAGCAAGTGTGATTTATAGAGTTATCAATCCATGGACAACT 566
 169 eileLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThr 186
 : : : : :
 567 TATTAATAAAACACTCCACAAAGTTTCAAGATCAAGCTTTAAACCTG 616
 186 alGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSerSerGly 202
 : : : : :
 617 TAGAAGTTAACGCAAGAACAAAGTTGCTCTTAAATGAACCTTTGGGT 666
 203 SerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLeuVal 219
 : : : : :
 667 TCACCGGTTTGGCAAAAGTGTGCAACTTACCATGAAAATTTGGTCTAC 716
 219 rArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValSerPro 236
 : : : : :
 717 GCGTTTCTTCACGCTAGAGATCCAAATTTATGAAAACCAAGTTTCCAC 766
 236 lYThrAlaIleLeuThrValValProPheHisHisGlyPheGlyMetPhe 252
 : : : : :
 767 GCACGGCTATTTTAACTAGTACCATTCATCATGCTTTGGTATGTTT 816
 253 ThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLeuThr 269
 : : : : :
 817 ACTACTTAAAGCTATCTAATCTGTGTGTTTCTATTTGTCATGTAAAGAA 866
 269 sPheAspGluGluThrPheLeuLysThrLeuLysAspTyrLysCysSer 286
 : : : : :
 867 ATTTCACGAGAGACTTTTAAACAACTGCAAGATTAACAATGTTCAA 916
 286 eValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
 : : : : :
 917 GCGTATCTCTGTACGACTTTGTTTGCATCTTATATGAAAGTAATTA 966

303 LeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyGlyAl 319
 : : : : :
 967 CTCGTAATATATGATTAATCAAAATTTAGTTGAAATTTGCATCTGGCGAGC 1016
 319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeu 336
 : : : : :
 1017 ACCTTATCTAAGAAATTTGCTGAAGCTGTGTCTAGACGTTTAAATTTAC 1066
 336 roGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAlaIle 352
 : : : : :
 1067 CGGCTGTCTGCAAGGCTATGTGTTTAAACAGAAACACTCTGCAATTAAT 1116
 353 IleThrProGluGlyAspAspLysProGlyValAspSerGlyLysVal 369
 : : : : :
 1117 ATCAACCGGAGCGCATGATTAACCAAGTCTTCTGCAAAAGTTGTGCC 1166
 369 oLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGly 386
 : : : : :
 1167 ATTATTTAAAGCAAAAGTTATCGATCTTGATTAATAAAAAAATTTGGGCC 1216
 386 roAsnArgArgGlyGlyValCysValLysGlyProMetLeuMetLysGly 402
 : : : : :
 1217 CGAACAGACGTCGAGAAAGTTGTGTTAAAGGCTCTATGCTTATGAAAGCT 1266
 403 TyrValAspAsnProGluAlaThrArgGluIleLeuAspGluGlyTyr 419
 : : : : :
 1267 TATGTAGATATATCCAGAACAGACAGAAATCATATGATGACAGAGTTG 1316
 419 pLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePhe 436
 : : : : :
 1317 GTTGACACAGAGAAATTTGGAATTTACGATGAAGAAACAAATTTCTTTA 1366
 436 leValAspArgLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
 : : : : :
 1367 TCAGGATCTGTTGAAGTCTTTATCAATCAAAAGATATCAAGTACCA 1416
 453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAsp 469
 : : : : :
 1417 CCGTGTGAATTAGAATCTGTTCTTTGCAACATCCAAATATTTTGTATGC 1466
 469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyVal 486
 : : : : :
 1467 CGCGCTGCTGGCGCTTCAGATCTTATAGCTGTGAGCTTCCGGAGCTG 1516
 486 alValValLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
 : : : : :
 1517 TTGTGTACTTGAAGAAAGAAATCTATGACTGAAAAAGAAATATGAGAT 1566
 503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValArg 519
 : : : : :
 1567 TACGTTGCTAGTCAAGTTTCAATGCAAAACGTTTCTGCTGTGTGCTCG 1616
 519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLys 536
 : : : : :
 1617 TTTTGTGAGAGAAAGTACCTAAGAGTCTCACTGTGTAATTTGACGATTAAG 1666
 536 lAlleArgGluIleLeuLysLysProValAlaLysMet 548
 : : : : :
 1667 CAATTAGAGAAATACTGAAGAAACAGTTGCTAAGATG 1704

seq_name: gb_pat: AR062708

seq_documentation_block: 1704 bp DNA

LOCUS AR062708 Sequence 5 from patent US 5843746. PAT 29-SEP-1999

DEFINITION AR062708

ACCESSION AR062708

VERSION AR062708.1 GI:5990399

KEYWORDS

SOURCE

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1704)

AUTHORS Tatsumi,H., Fukuda,S., Kikuchi,M. and Koyama,Y.

TITLE

Biotinlated firefly luciferase, a gene for biotinated firefly luciferase, a recombinant DNA, a process for producing biotinated luciferase and a bioluminescent analysis method

JOURNAL

Patent: US 5843746-A 5 01-DEC-1998;

FEATURES

Location/Qualifiers
1..1704

Source
/organism="unknown"

BASE COUNT 541 a 275 c 361 g 527 t
ORIGIN

alignment_scores:

Quality: 2801.00 Length: 546
Ratio: 5.130 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.451

alignment_block:

US-09-581-241-4 x AR062708 ..

Align seg 1/1 to: AR062708 from: 1 to: 1704

```

3 AsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrPr 19
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
67 AGTCGAGACGATGATAATATTGATGATGATGATGATGATGATGATGATGAT 116
19 otleGluGluGlySerAlaGlyAlaGluIleuArgLysTyrMetAspArg 36
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
117 TATTGAAGGGGATCTGAGACACAAATGGCGCAAGTATATGATGATGATGAT 166
36 TyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValAsp 52
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
167 ATCCAAACTTGGAGCAATGCTTTACTACGCACTTACCGGTTGCGAT 216
53 TyrThrTyrAlaGluTyrLeuGluLysSerCysGlyGlyGluAlaIle 69
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
217 TATACGTACGCCGAAATCTAGAAAAATCATGCTCTCTAGAGAGGCTTT 266
69 uLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCysSerGlu 86
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
267 AAGAAATTAAGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTT 316
86 snCysGluGluPhePheIleProValIleuAlaGlyLeuPheIleGlyVal 102
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
317 ACGTGAGAGGCTTCTATCTCTGATTAAGCGGTTTATTTATGATGATGAT 366
103 GlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHis 119
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
367 GGTGTGCTCCAACTAATGATTTACACTCTACGTGAATGGTTCACAG 416
119 rLeuGlyIleSerLysProThrIleValIlePheSerSerLysGlyLeu 136
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
417 TTTAGGATCTTAAGCCACAAATTTGATTTAGTTCTAAAAAAGGATTA 466
136 sPlyValIleThrValGlnLysThrValThrAlaIleLysThrIleVal 152
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
467 ATAAGTTAATCTTACAAAAAAGGTTAACTCTATTAATAAACCATTTG 516
153 IleLeuSerLysValAspTyrArgGlyTyrGlnSerMetAspAsnPh 169
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
517 ATATTGACGCAAGTGAATTAAGAGTTATCAATCCATGGACAACTT 566
169 eIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThr 186
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
567 TATTAAAAAAACACCTCCACAGTTTCAAGATCAAGTTTAAAACTG 616
186 aGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSerSerGly 202
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
617 TAGAAGTTAACCGCAAGAACCACTGCTCTTAATGAACCTTCGGGT 666
203 SerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLeuVal 219
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
667 TCACCGGTTGGCCAAAAGGTGTGCAACTTACTCATGAAAAATTTGGTCAC 716

```

```

219 rArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValSerProG 236
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
717 GCGTTTTTTCACCTTGACATCCAAATTTATGAAACCAAGTTTCACAG 766
236 TyrThrAlaIleLeuThrValValProPheHisHisGlyPheGlyMetPhe 252
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
767 GCACGGCATTTTAACTGATGATACATTCATCATGATGATGATGATGAT 816
253 ThrThrLeuGlyTyrLeuThrThrCysGlyPheArgIleValMetLeuThr 269
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
817 ACTACTTAGGCTATCTACCTGTTGGTTGGTTGGTTGGTTGGTTGGTT 866
269 sPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSer 286
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
867 ATTGACCAAGACCTTTTAAAAACACGCAAGATTAACAAATGTTCAA 916
286 eValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
917 GCGTTATCTTGTACCGCCTTTGTTGCAATCTTAATAGAAAGTAA 966
303 LeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyGlyAl 319
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
967 CTGATTAATATGATTTATCAAAATTTAGTGAATTCATCTGGCGGAGC 1016
319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeu 336
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1017 ACCTTATCTAAGAAATTTGGTGAAGCTGTTGTAAGCTTTTAAATTAC 1066
336 rGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAlaIleIle 352
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1067 GCGGTGTCGTCACAGCTATGTTGTTAAACAAACACTCTGCAATATT 1116
353 IleThrProGluGlyAspAspLysProGlyAlaSerGlyLysValAla 369
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1117 ATCACACCGCAAGCGGATGATTAACACAGTCTCTGCAAAAGTTGGCC 1166
369 oLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGly 386
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1167 ATTATTTAAAGCAAAAGTTATCGATCTTGATCACTAAAAAACTTTGG 1216
386 rAsnArgArgGlyGluValCysValLysGlyProMetLeuMetLysGly 402
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1217 CCAACAGACGTGAGAAAGTTGTGAAGGTCCTATCTTATGAAAGT 1266
403 TyrValAspAsnProGluAlaThrArgGluIleIleAspGluGlyTyr 419
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1267 TATGTAGATTAATCCAAAGCAACAGAAATCATATGAAAGAGTTG 1316
419 pLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePhe 436
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1317 GTTGACACAGAGATATTGGGATTTACGATGAAGAAAAAACAATTTCT 1366
436 lValAlaPargLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1367 TCGTGATCTGTTGAAGTCTTTATCAAAATTAACAAAGATATCAATACA 1416
453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAsp 469
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1417 CCGTCTGAATTGAATCTGTTCTTTGCAACATCCAAATATTTTGAATG 1466
469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyAla 486
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1467 GCGGTTGGTGGGCTCCAGATCTTATACCTGATGAGCTTCGGGAGCTG 1516
486 aValValLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1517 TTTTGTACTTGAAAAAGAAATCATATGACTGAAGAAAAAGATATGAT 1566
503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyGlyVal 519
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1567 TACGTGTGCTGTCMAAGTTCAAAATGCAAAACGTTTCGGGTGCTCG 1616
519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLys 536

```



```

|||||
1617 TTTTGTGCGAGTAACCTPAAGTCTCAGTGAATAATTGACGGTAAG 1666
536 lallearglulileuylslyspovalalalysmet 548
1667 CAATTAGAGAAATFACTGAGAAACAGTTGCTAAGATG 1704
seq_name: gb-pat:E12278

seq_documentation_block:
LOCUS      E12278      1704 bp      DNA      PAT      24-JUN-1998
DEFINITION CDNA encoding biotinylated Luciola luciferase.
ACCESSION  E12278
VERSION    E12278.1 GI:3251112
KEYWORDS   JP 1996308578-A/3.
SOURCE     Luciola lateralis.
ORGANISM   Luciola lateralis
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
            Elateriformia; Cantharoidae; Lampyridae; Luciola.
REFERENCE  1 (bases 1 to 1704)
            Tatsumi, H., Fukuda, M., Kikuchi, M. and Koyama, T.
            BIOTIN FIRE FLY LUCIFERASE, BIOTIN FIRE FLY LUCIFERASE GENE, NEW
            RECOMBINANT DNA, PRODUCTION OF BIOTIN FIRE FLY LUCIFERASE AND
            BIOLUMINESCENCE ANALYSIS
            Patent: JP 1996308578-A 3 26-NOV-1996;
            KIKOMAN CORP
COMMENT     OS      Luciola lateralis
            PN      JP 1996308578-A/3
            PD      26-NOV-1996
            PE      24-APR-1995 JP 1995098857
            PR      27-JUL-1994 JP 94P 193798, 14-MAR-1995 JP 95P 54625 PI
            TS      TATSUMI HIROKI, FUKUDA MASARU, KIKUCHI MAMORU, KOYAMA TADUO
            CI      C12N15/09,C12N9/02,C12O1/26,C12N9/02,C12R1/185; CC
            ST      Strandness: Double;
            CC      topology: Unknown;
            CC      hypothetical: NO;
            CC      anti-sense: NO;
            FH      Key      Location/Qualifiers
            FT      source      1..1704
            FT      mat_peptide 1..1704
            FT      /c1one='pHLf203',/organism='Luciola lateralis' FT
            FT      /product='biotinylated Luciola luciferase'.
FEATURES
     source      1..1704
                  Location/Qualifiers
     BASE COUNT  541 a      275 c      361 g      527 t
     ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

167 ATGCAAAACTTGAGCAATTGCTTTACTACCGCACTTACCGGTGCGAT 216
53 TTTTThTyrAlaGluTyrLeuGluIuysSerCysLeuGluValAla 69
217 TATAGTACGCCGCAATFACTAGAAAAATCATCTGCTGAGAGAGCTT 266
69 uLysAsnTyrGlyLeuValAlaSpGlyArgGluLeuAlaLeuCysSerGlu 86
267 AAGAAATTTATGCTTGGTTGGTTGATGAGAGAAATTCGTTATCCAGTGA 316
86 snCysGluGluUphePheIleProValAlaGluValLeuPheIleGlyVal 102
317 ACTGTGAAGAGTCTTTATTCCTGATTAAGCCGGTTATTTATAGTGTG 366
103 GlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHis 119
367 GGTGGCTCCAACTAATGAGATTACACTGCTGATGATGTTGTTGACAG 416
119 rLeuGlyIleSerLysProThrIleValPheSerSerLysIleGlyLeu 136
417 TTTAGGCATCTCTAAGCCACAACTTGTATTTAGTTCTAAAAAGGATTA 466
136 sPLysValIleThrValGlnIleThrValThrAlaIleLysThrIleVal 152
467 ATAAAGTTATTAAGTGTACAAAAAAGGTAAGTGTATTAACCATTTGTT 516
153 rLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAspAsn 169
517 ATATTGACACAAAGTGATTAATGAGTTATCAATCAAGCAAGCACTT 566
169 eLelysLysAsnThrProGlnIlePheLysGlySerSerPheLysThr 186
567 TATTAATAAAAAAACACTCCACAGGTTCAAGGATCAAGTTTAAAGTG 616
186 alGluValAsnArgLysGluGluValAlaLeuIleMetAsnSerSerGly 202
617 TAGAAGTTAACGCCAAAGAACAAAGTGTCTTATTAATCAACTCTTCGGGT 666
203 SerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnLeuVal 219
667 TCACCGGTTTGGCAAAAGTGTCACACTTCTCATGAAATTTGGTCAC 716
219 rArgPheSerHisAlaArgAspProIleTyrGlyAsnGluValSerPro 236
717 GCGTTTTTCTCACGCTAGAGATCCAAATTATGAAACCAAGTTTCACAG 766
236 lYThrAlaIleLeuThrValValProPheHisHisGlyPheGlyMetPhe 252
767 GCAGCGCTATTTTACGTGTAGTACCATTCATCATGTTGGTATGTT 816
253 ThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLeuThr 269
817 ACTACTTTAGGCTATCTAAGTGTGGTTTCGTATTTCTCATGTTAAGAA 866
269 sPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSer 286
867 ATTTGACGAAGAGACTTTTAAAAACATGCAAGATTAACAAATGTCAA 916
286 eValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu 302
917 GCGTTATCTTGTACCGCACTTGTTCATTTCTTAATTAAGATGAATTA 966
303 LeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyVal 319
967 CTCGATAAATATATTTATCAAAATTTAGTGAATGTCATCTGCGGAGC 1016
319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeu 336
1017 ACCTTATCTTAAGAAATGTTGTAACCTGTTCTAAGCTTTTAATTTAC 1066
336 rGlyValArgGlnGlyTyrGlyLeuThrGluThrHisSerAlaIle 352
1067 CGGGTGTGCTCAAGGCTATGTTTAACAGAAACAACCTTGCATTAAT 1116

```



```

353 11ethrProGluGlyAspAspLysProGlyAlaSerGlyLysValValPr 369
|||||
1117 ATCACACCGGAGGCGATGATTAACACAGTGCTGCTGCGCAAGATTGTGCC 1166
369 OleuThelysAlaLysValIleAspLeuAspThrLysThrLeuGlyP 386
|||||
1167 ATTATTATTAAGCAAAAGTTATCGATCTTGATCTAAATAAAACCTTTGGGCC 1216
386 roAsnArqrGlyGlyLysValCysValLysGlyProMetLeuMetLysGly 402
|||||
1217 CGAACGACAGCTGAGAGATTTGCTTAAGGGCTCATCTATTATGAAGGT 1266
403 TyrValAspAsnProGluAlaThrArgGluIleIleAspGluGlyTr 419
|||||
1267 TATGTGATTAATCCAGAACCAAGCAAGCAATCATGATGAAGAGGTGG 1316
419 pleuHsThrGlyAspIleGlyTyrTyrAspGluGlyLysHisPhePhe 436
|||||
1317 GTTGCAACAGAGATATTGGTATTACATCAAGAAAAACATTTCTTTA 1366
436 1LeValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGluValPro 452
|||||
1367 TCGTGATGCTTGAAGCTCTTAATCAATACAAAGATATCAAGTACCA 1416
453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAl 469
|||||
1417 CCTGCTGAATTAAGAACCTGTTCTTTGCAACATCCAAATATTTTGTATCC 1466
469 aglyValAlaGlyValProAspProIleAlaGlyLysLeuProGlyAlaVal 486
|||||
1467 CGCGCTGCTGCGGCTCCAGATCCTTACTGCTGAGCTTCCGGAGCGTG 1516
486 a1ValValLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
|||||
1517 TTGTGTACTTGAAGAAAGCAAAATCTATGACTGAAAAAGAGTAATGAT 1566
503 TyrValAlaSerGluValSerAsnAlaLysArgLeuArgLysValAlaArg 519
|||||
1567 TACGTGCTAGTCAGATTTCATTAATGCAAAACGTTTCGCTGCTGCTCCG 1616
519 gpheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysAs 536
|||||
1617 TTTTGTGAGAGAGTACCTTAAGGCTCTCACTGCTAAATGACGTAAG 1666
536 1a1leArGluIleLeuLysLysProValAlaLysMet 548
|||||
1667 CAATTAGCAAAATACGAAAGAACCAAGTTGCTAAGATG 1704
seq_name: gb_pat:E13411
seq_documentation_block: 1704 bp DNA PAT 24-JUN-1998
LOCUS E13411
DEFINITION cDNA encoding biotinylated firefly luciferase bl203.
ACCESSION E13411
VERSION E13411.1 GI:3252216
KEYWORDS JP 1997187281-A/2.
SOURCE unidentifed.
ORGANISM unidentifed.
REFERENCE 1 (bases 1 to 1704)
AUTHORS Masuda,S., Tatsumi,H. and Koyama,T.
TITLE ANTIBODY-FIREFLY LUCIFERASE FUSION PROTEIN, GENE OBTAINED THEREFROM, NOVEL RECOMBINANT DNA AND PRODUCTION OF ANTIBODY-FIREFLY LUCIFERASE FUSION PROTEIN
JOURNAL Patent: JP 1997187281-A 2 22-JUL-1997;
COMMENT OS Luciola lateralis (firefly)
KIRKMAN CORP
PN JP 1997187281-A/2
PD 22-JUL-1997
PR 09-JAN-1996 JP 1996001812
PI MASUDA SUSUMU, TATSUMI HIROKI, KOYAMA TAJIJI
PC C12N15/09,A61K39/395,C07H21/04,C07K14/47,C07K19/00,C12P21/02,

```

```

PC G01N33/535//
PC C12N9/02,(C12P21/02,C12R1:19);
CC strandedness: Double;
CC topology: Unknown;
CC hypothetical: NO;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..1704 /organism='Luciola lateralis' FT
FT /clone='pHLF200' /product='Biotinylated firefly luciferase FT
FT mat_peptide 1..1704
FT /db_xref='taxon:32644'
FEATURES
source location/Qualifiers
1..1704 /organism='unidentifed'
/db_xref='taxon:32644'
BASE COUNT 541 a 275 c 361 g 527 t
ORIGIN
alignment_scores:
Quality: 2801.00 Length: 546
Ratio: 5.130 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.451
alignment_block:
US-09-581-241-4 x E13411 ..
Align seg 1/1 to: E13411 from: 1 to: 1704
3 AsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPheTyrPr 19
|||||
67 AGCTCGAAGACGATGAAATATGTATGCTCGTGAACCATTTTACCC 116
19 o1leGluGlySerAlaGlyAlaGlnLeuArgLysTyrMetAspArgT 36
|||||
117 TATGAGAGGAGATCTGCTGAGCAATTCGCGCAAGATATGATGAT 166
36 TyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValAsp 52
|||||
167 ATGCNAACCTTGACCAATGCTTTACTACGCACTTACCGGTGTGAT 216
53 TyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGlyGluAlaLe 69
|||||
217 TATAGCTAGCCGCAATTAATTAGAAAAATCATGCTGTAGAGAGGCTTT 266
69 uLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuLysSerGlu 86
|||||
267 AAGCAATTAATGTTGTTGTTGATGAGAGAAATTCGTTATGCACTGAAA 316
86 snCysGluGluPhePheIleProValIleuAlaGlyLeuPheIleGlyVal 102
|||||
317 ACTGTGAAGAGTTCTTATCTCTGATTTAGCCGCTTATTTATAGTGTC 366
103 GlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuValHis 119
|||||
367 GGTGTGCTCCAACTAATGAGATTACACTTACGTAATGTTGTTACAG 416
119 rLeuGlyIleSerLysProThrIleValPheSerSerLysGlyLeuAl 136
|||||
417 TTTTGCATCTCTTAACCAACAACTTATTTAGTTTAAAGAGATTAG 466
136 sPlsValIleThrValGlnLysThrValThrAlaIleLysThrIleVal 152
|||||
467 ATTAAGTTTAATTAAGTACAAAAAGCTAAGCTATTAAACCATTTGT 516
153 IleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAspAsp 169
|||||
517 ATATTGACAGCAAGATGATTAATAGAGTTTAAATCAATCATGACAACT 566
169 e1leLysLysAsnThrProGlnIlePheLysGlySerSerPheLysThrV 186

```



```

alignment_scores:
  Quality: 2801.00      Length: 546
  Ratio: 5.130         Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.451

```

```
alignment_block:
  uc-00-581-241-4 v F10169
```

Align seg 1/1 to: E10169 from: 1 to: 2019

3 AsMerglAsnAspGluAsnIleValIleTYrGlyProGluProPheTYrPr 19
 382 TCCCTGAGACACATGAAATATTTGTTGTTATGTCCTGACACATTTTACC 431
 19 oIleGluGluGlySerAlaGlyAlaGluLeuArgLysTYrMetAspArgT 36
 432 TATTGAAGAGCGATCTGCTGGACACACATTTGCCCAAGTAATATGATCAT 481
 36 TyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGlyValAsp 52
 482 ATGCAAACTTGAAGCATTTGCTTTTACTAACGGCACTTACCGGTGTGAT 531
 53 TYrThrTYrAlaGluTYrLeuGluLysSerCysCysLeuGlyGluAlaLe 69
 532 TATACGTACGCCCAATACCTTAGAAAAATCATGCTGTACGAGAGCGCTT 581
 69 uLysAsnTYrGlyLeuValValAspGlyAlaGlyIleAlaLeuCysSerGlu 86
 582 AAGCAATTAAGTTGGTTGGTGTGTGATGAGAGAAATTCGGTTATGCAAGTAA 631
 86 snCysGluGluPhePheIleProValLeuAlaGlyLeuPheIleGlyVal 102
 632 ACGTGAAAGGTCTTTATTCCTGTATATAGCCGGTTATTATATAGTGTTC 681
 103 GlyValAlaProThrAsnGluIleTYrThrLeuArgGluLeuValHisSe 119
 682 GGTGTGGCTCCCACTAATGAGATTATACCTACGTGGAATTTGGTTCACAG 731
 119 rLeuGlyIleSerLysProThrIleValPheSerSerLysLysGlyLeuAl 136
 732 TTTAGGCATCTCTAAGCCAAATTTGATTAGTTCTAAAAAAGATATAG 781
 136 sPLysValIleThrValGlnLysThrValThrAlaIleLysThrIleValAl 152
 782 ATAAAGTATTAACGTACAAAAAAACGGTAACTGCTATTAATAAACCATTC 831
 153 IleLeuAspSerLysValAspTYrArgGlyTYrGlnSerMetAspAspPh 169
 832 ATATTTGGACACCAAGTGGATTATAGAAGCTTATCATATGCATGGACAAC 881
 169 eIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThrV 186
 882 TATTAAAAAAACCTCCACAAGGTTTCAAAAGATCAAGTTTAAACCTG 931
 186 alGluValAsnArgLysGluGlnValAlaIleuIleMetAsnSerSerGly 202
 932 TAGAAGTTAAACCGCAAGACAAAGTGGCTTATATATGAACTCTTCGGGT 981
 982 TCAACCGGTTGGCCAAAAGTGTGCACACTCTCATGAATAATTTGGTCC 103
 219 rArgPheSerHisAlaArgAspProIleTYrGlyLysnGlnValSerProG 236
 1032 GCGTTTTCCTACCGCTAGAGATCCCAATTTATGGAACCAAGTTTCAACG 108
 236 IYrThrAlaIleLeuThrValValProPheHisnIleGlyPheGlyMetPhe 252
 1082 GCACGGCATTTTAATAGTATGACATTCGCAATTCAGATGGTTTGGTATG 113
 253 ThrThrLeuGlyTYrLeuThrCysGlyPheArgGlyIleValMetLeuThrLy 269

1132	ACTACTTTAAGGCTATCTAACTGCTGGTTTGGTATGTGTCATGTTAAAGCAA	1181
269	sPheaspGIugluThrPheLeuLysThrLeuGlnAspTyrTrpCysSers	286
1182	ATTGTCGAAGAGACTTTTAAAAACATGCACAAATTACAAAAGTTCAA	1231
286	erValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGluLeu	302
1232	GGCTTATCTCTGTACCGCACTTGTTGGCATCTTAATAGAAGGAATTA	1281
303	LeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyGlyAl	319
1282	CTCGATTAATATGATTTATCAAAATTAGTGAATATGCATCTGGGGAGC	1331
319	apRhoLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeuP	336
1332	ACCTTATCTAAAGAAATTGGTGAACCTTCTAGACGTTTAATTATTC	1381
336	roGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAlaIleIle	352
1382	CGGCTGTCTCAAGGCTATGGTTTAAACAGAAACAACCTGCAATTTAT	1431
353	IleThrProGlnGlyAspAspLysProGlyAlaSerGlyLysValValPr	369
1432	ATCACACCGGAAGCGCATGTAAACAGAGTCTTTCGCAAAAGTTGTGC	1481
369	oleuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP	386
1482	ATTATTTAAAGCAAAAGTTATGCATCTTATATCTAAAAAACTTTGGGC	1531
386	roAsnArgArgGlyGluValCysValLysGlyProMetLeuMetLysGly	402
1532	CGAACACACGTGGAGAGATTGTGTAAAGGTCCTATGCTTATGAAGGT	1581
403	TyrValAspAsnProGluAlaThrArgGluIleIleAspGluGluGlyTr	419
1582	TATGTATATATCAACAGACAAACAGAAATCATGATGAAGAAGTTG	1631
419	pLeuHisThrGlyAspIleLeuTyrTyrAspGluGluLysHisPhePheI	436
1632	GTTCACACAGGAGATTTGGTATTACAGTGAAGAAAAACATTTCTTAA	1681
436	IleValAspArgLeuLysSerLeuIleLysTyrLysGlyGluLeuProG	452
1682	TCGTGGATCGTTTGAATCTTTATCAATCAAGAAAGATATCAAGTACCA	1731
453	ProAlaGluLeuGlnSerValLeuLeuGlnHisProAsnIlePheAspAl	469
1732	CTGCTCAATATTAGATTCGTTCTTTGGCAACATCAAAATATTTTGATAC	1781
469	agIValAlaGlyValProAspProIleAlaGlyGluLeuProGlyAlaV	486
1782	CGGCGTTCCTGGCTTCACATCTCTATAGCTGGTGAAGCTTCGGGAGCTG	1831
486	alValAlaLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp	502
1832	TTGTGTACTTAAAAAGCAAAATCTATGACTCAAAAAAGAAAGTAAAGAT	1881
503	TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValAlaAr	519
1882	TACGTGTCTACTCAAGTTTCAATGCAAAACGTTTGGCGTGTGTGTCCG	1931
519	qPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysA	536
1932	TTTTGTGACAGCAATACTTAAGGTCTCACTGGTAAATATGACGATAAAG	1981
536	IaIleArgGluIleLeuLysLysProValAlaLysMet	548
1982	CAATTAGGAATACTGAAGAAACCAAGTCTCTAAGATG	2019


```

84  erGluSnCySGluGluPhePheIleProValIleuAlaGluLeuPheIle 100
    |||
251  GTGAAACCTGTGAAGATCTTATTCCTGATATAGCGGTTATTTATA 300
    |||
101  GLVAlGluValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVa 117
    |||
301  GGTGTGGTGTGGCTCCAACTAATGAGATTACACTGCTACGTGAATGGT 350
    |||
117  IHisSerLeuGlyIleSerLysProThrIleValIlePheSerLysLysG 134
    |||
351  TCACAGTTAGGATCTCTAAGCCACAACATTTGATTAGTTCAAAAAAG 400
    |||
134  ILeuAspLysValIleIleThrValGluLysThrValIleAlaIleLys 150
    |||
401  GATTGATTAAGTTATATACGTACAAAAACGGTAACGCTATTAAAAACC 450
    |||
151  ILeValIleLeuAspSerLysValAspIleArgGlyTyrGlnSerMetAs 167
    |||
451  ATTGTTATATTTGACACGACCAAGTGATATAGAGTTATCAATCAATGGA 500
    |||
167  PAsnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
    |||
501  CAACCTTATTTAAAAAACAACCTCCACAAGGTTCAAGATCAAGTTTAA 550
    |||
184  YAspThrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
    |||
551  AAACGTGTAGAGTTAAACCGCAAGAACAAAGTTGCTCTTATTAAGAAGCT 600
    |||
201  SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnI 217
    |||
601  TCGGGTTCACCGGTTGGCAAAAGGTGTGCACATCTACATCAAAAAATAT 650
    |||
217  eValIleThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
    |||
651  CGTCACTAATTTTTCACGTAGAGATCCAAATTATGGAACCAAGTTT 700
    |||
234  ePProGlyThrAlaIleLeuThrValIleProPheHisIleGlyPheGly 250
    |||
701  CACCGAGCGCGCTATTTTAACTGATGACCATTCATCATGTTTGGT 750
    |||
251  MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
    |||
751  ATGTTTACTACTTTAGGCTATCTAAGCTTGTGGTTTCGATTTGTCATGTT 800
    |||
267  UThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
    |||
801  AACGAAATTTGACGAGAGACACTTTTAAAAACACTGCAAGATTCAAAAT 850
    |||
284  YSerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
    |||
851  GTTCAAGGTTATTTCTGTACCGACTTTTGTGCATTTCTTAATAGAAAGT 900
    |||
301  GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
    |||
901  GAATTAACCTCGATTAATGATTTATCAAAATTAGTTGAATGCAATCTGG 950
    |||
317  YGlyValAlaProLeuSerLysGluIleGlyGluValAlaAlaArgArgPhe 334
    |||
951  CGGACACCTTTATCTAAAGAAATGTTGAAACGCTGTGCTAGACGTTTAA 1000
    |||
334  snLeuProGlyValAlaArgGlnGlyTyrGlyLeuThrGluThrSerAla 350
    |||
1001  ATTTAACCGGGTGTGTCAGGCTATGTTTAAACGAAAAACACTCTGCA 1050
    |||
351  IleIleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
    |||
1051  ATTATTTATCACACCGAAGCGGATGATTAACCAAGTGCTTGCAGCAAGT 1100
    |||
367  IValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysIle 384
    |||
1101  TGTGCAATTTATTTAAAGCAAAAGTTATCGATCTTCACTAAAAAACCTT 1150
    |||
384  euGlyProAsnArgArgGlyGluValLysValLysGlyProMetLeuMet 400

```

```

1151  TGGGCCCGACAGACGCTGAGAAAGTTTGTGTAAGAGGTCCTATGCTTATG 1200
    |||
401  LysGlyTyrValAlaAspAsnProGluAlaThrArgGluIleIleAspGlu 417
    |||
1201  AAAGTTATCTAGATTAATCCAGAACCAAGAAAGAAATCATAGATGAAGA 1250
    |||
417  uGlyThrPheHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
    |||
1251  AGGTTGGTTCACACAGAGAGATATGGGTATTCAGATGAAGAAAAACATT 1300
    |||
434  hePheIleValAlaAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
    |||
1301  TCTTTATCGTGATCGTTTGAAGCTTTATCAATAACAAAGATATACAA 1350
    |||
451  ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIleP 467
    |||
1351  GTACCACTGCTGTAATTAGAATCTGTTCTTTTCACACATCCAAATATTTT 1400
    |||
467  eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
    |||
1401  TGATGCCGGGCTGTGCTGGCTTCCAGATCTTATAGCTGGTGAGCTTCGG 1450
    |||
484  IValAlaValAlaValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
    |||
1451  GACCTGTTGTTGACTTAAGAAAGAAATCTATGACTGAAAAAGAAATA 1500
    |||
501  MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
    |||
1501  ATGGATTACGTTCTAGTCAAGTTTCAAAATGCAAAACGTTGGTGCTGG 1550
    |||
517  YValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
    |||
1551  TGTCCGTTTGTGGAGCAAGTACCTTAAGGTCTCACTGGTAAATTTGAGC 1600
    |||
534  IYlyAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
    |||
1601  GTAAGCAATTAGAGAAATAGCAAGAAACAGTTGCTAAGATG 1644
    |||
seq_name: gb_pat:E36479
seq_documentation_block: 1644 bp DNA PAT 07-FEB-2001
LOCUS E36479
DEFINITION Luciferase and method for assaying intracellular ATP by using the
same.
ACCESSION E36479
VERSION E36479.1 GI:13022686
KEYWORDS JP 1999239493-A/3.
SOURCE Luciola lateralis.
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
Elateriformia; Cantharoidae; Lampyridae; Luciola.
REFERENCE Noriaki,H.S.M.M.
AUTHORS
TITLE Luciferase and method for assaying intracellular ATP by using the
JOURNAL Patent: JP 1999239493-A 3 07-SEP-1999;
KIKKOMAN CORP
COMMENT OS Luciola lateralis
PN JP 1999239493-A/3
PD 07-SEP-1999
PF 21-DEC-1998 JP 1998363108
PR
PI NORIAKI HATTORI,SEIJI MURAKAMI
PC C12N15/09,C12N9/02,C12Q1/66/(C12N9/02,C12R1:19),C12N15/00 CC
FEATURES
source location/Qualifiers
FT key CDS (1)..(1644).
1..1644 /organism="Luciola lateralis"
/db_xref="taxon:7052"
BASE COUNT 528 a 261 c 349 g 506 t

```

ORIGIN

alignment_scores:
 quality: 2821.00 length: 548
 ratio: 5.148 gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.818

alignment_block:
 US-09-581-241-6 x E36479

Align seg 1/1 to: E36479 from: 1 to: 1644

```

1 MetGluAsnMetGluAsnAspGluAsnIleValIyrGlyProGluProph 17
1 ATGGAAAACATGAGAACGATGAAAATATGTGTATGCTCTGACCAATT 50
17 eTyPProIleGluGluGlySerAlaGlyAlaGluLeuArgIlyStyMet 34
51 TTACCCCTTTGAAAGGAGATCTGCTGGAGCACATTTGGCAAGTATATGC 100
34 sPaRgTyAlaIylsleuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 ATCGATATGCAAAACTTGACCAATTCCTTTACTAACGCACTTACCGGT 150
51 ValAspTyThrTyAlaGluTyIleuGluIylsSerCysIleuGlyGly 67
151 GTCGATTTACGTACGCCGAATCTTAGAAAATCATCTGCTCTAGAGGA 200
67 uAlaLeuIylsAsnTyGlyIleuValAlaIylsGlyArgIleAlaIleu 84
201 GGCTTTAAAGAAATTATGTTGGTTGGTTGATGGAAGAATTCGGTATGCA 250
84 eRgIuAsnCysGluGluPhePheIleProValIleuAlaGlyLeuPheIle 100
251 GTGAAGAACTGTCAGAAATCTTATCTGATTTAGCCGGTTATATTATTA 300
101 GylValIylsAlaIylsProThrAsnGluIleTyThrLeuArgGluLeu 117
301 GGTCGCGGTGGCTCCAACTAATAGATTTACCTACGTGGAATTCGT 350
117 hIleSerIleuGlyIleSerIylsProThrIleValPheSerSerIyls 134
351 TCACAGTTTAGCATCTCTAAAGCAAACTGATTTAGTCTTAAAGAAAG 400
134 IylsLeuAspIylsValIleThrValIylsIylsThrValIleAlaIle 150
401 GATTAGATTAAGTTATTAAGTACAAAACGCTACGCTATTTAAACC 450
151 IleValIleLeuAspSerIylsValAspTyArgGlyTyArgIleuSer 167
451 ATTGTTATTTGGACAGCAAAAGTGATTTAGAGTTATCATCATGATGA 500
167 pAsnPheIleIylsIylsAsnThrProGlnIylsPheIylsSerPheIle 184
501 CAACCTTATTTAAAAAACAACCTCCACAGGTTTCAAGGATCAAGATTT 550
184 yAsThrValIylsValAsnArgIylsGluIylsAlaIleuIleMetAsn 200
551 AAACGTGTAAGTTAACCGCAAAAGAACAGTTCCTTTATTAATCAACT 600
201 SerGlySerThrGlyLeuProIylsGlyValAlaIleuThrHisGluAsn 217
601 TCGGGTTCAACCGGTTTGGCAAAAGGTGCAAACTTACTCATGAATAAT 650
217 eValIThrArgPheSerHisAlaArgAspProIleTyArgIylsAsnGlu 234
651 GGTCACTAATTTTCTACGCTAGAGATCAATTTATGGAACCAAGTTT 700
234 ePrGlyThrAlaIleuThrValIylsProPheHisIylsGlyPheGly 250
701 CACCAAGCAGCGCTATTTTAACTGTAAGTACATTCATCATGTTTGGCT 750

```

```

251 MetPheThrThrLeuGlyTyThrLeuThrCysGlyPheArgIleValMetLe 267
751 ATCTTACTACTTTAGCTATCTAACTTGTGTTTGGTATGTCATGTCATGTT 800
267 uThrIylsPheAspGluGluThrPheLeuIylsThrLeuGlnAspTyIyls 284
801 AACGAATTTGACGAGAGACCTTTTAAAAACACTGCAAGATTCACAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleuAsnArgSer 300
851 GTTCAAGCGTTATCTTGTACGACCTTTGTTGCAATTTCTTANTAGAACT 900
301 GluLeuLeuAspIylsTyArgIylsPheSerAsnLeuValGluIleAlaSer 317
901 GAATTTACTGATTAATATATGATTTATCAATTTAGTTGAATTCATCTGG 950
317 yGlyAlaIylsProLeuSerIylsGluIleGlyValAlaValAlaArgArg 334
951 CCGAGCACCTTTATCTAAAGAAATGCTGAACCTGTTGCTAGACGTTTAA 1000
334 snLeuProGlyValArgGlnGlyTyArgIylsLeuThrGluThrThrSer 350
1001 ATTACCGGCTGTTTCGCAAGGCTATGTTTACAGAAACCACTCTGCA 1050
351 IleIleIleThrProGluGlyAspAspIylsProGlyAlaSerGlyIyls 367
1051 ATTATTTACACCGCAAGCGATGATTAACCAAGGCTCTCTGCAAAAG 1100
367 ValProLeuPheIylsAlaIylsValIleAspIleAspThrIylsIyls 384
1101 TGTGCCATTTATTAAGCAAAAGTTATCGATCTGTGATCTAAAAAAACT 1150
384 euGlyProAsnArgArgGlyIylsValIylsValIylsGlyProMetLeu 400
1151 TGGCCCGCAACGACGCTGAGAAAGTTGTGTAAGGCTCTATGCTTATG 1200
401 LysGlyTyThrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
1201 AAAGTTATGTAGATTAATCCAGAAACCAAGAAATCATATGATGAGA 1250
417 uGlyTyThrLeuHisThrGlyAspIleGlyTyTyArgGlyIylsIylsIsp 434
1251 AGGTTGGTGCACACAGAGATTTGGGCTATTTACGATGAAGAAAAACATT 1300
434 hePheIleValAspArgLeuLeuSerLeuIleIylsTyIylsGlyTyGln 450
1301 TCTTATCGTGCATCGTTTGAAGCTTTAATCAAAATCAAAAGATATCAA 1350
451 ValProProAlaGluLeuGluIylsValLeuGlnHisProAsnIlePhe 467
1351 GTACCACCTGCTGAATTAAGATTCGTCTTGTGCAACATCCAAATATTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyIylsLeuPro 484
1401 TGATGCCGCGCTTGTGCTGGCTCCAGATCCATAGCTGGGAGCTTCGG 1450
484 IylsAlaValAlaIylsLeuIylsGlyIylsSerMetThrGluIylsGlu 500
1451 GAGCTGTTGTTGTTACTTAAGAAAGAAATCTATGACTGAAAAAGAGTA 1500
501 MetAspTyThrAlaIylsSerGlnValSerAsnAlaIylsArgLeuArg 517
1501 ATGATTAACGTTGCTAGTCAAGTTTCAAAATGCCAAACCTTTGCTGGT 1550
517 yValArgPheValAspGluValProIylsGlyLeuThrGlyIylsIleAsp 534
1551 TGTCCGTTTGTGACGAGATCACTAAAGTCTCACGTGTAATTAATTAAG 1600
534 IylsAlaIleArgGluIylsLeuIylsIylsProValAlaIylsMet 548
1601 GTAAAGCAATTAAGAAATTAAGTGAAGAAACCAAGTTGCTAAGATG 1644

```

seq_name: gb_pat:AR098258

seq_documentation_block:

LOCUS AR098258 1644 bp DNA

DEFINITION Sequence 13 from patent US 6074859.

PAT

14-FEB-2001

ACCESSION AR098258

VERSION AR098258.1 GI:12807515

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1644)

AUTHORS Hirokawa,K., Kajiyama,N. and Murakami,S.

TITLE Mutant-type bioluminescent protein, and process for producing the

JOURNAL mutant-type bioluminescent protein

FEATURES Patent: US 6074859-A 13 13-JUN-2000;

Location/Qualifiers

1..1644

BASE COUNT 528 a 260 c 350 g 506 t

ORIGIN

alignment_scores:

Quality: 2815.00

Ratio: 5.146

Percent Similarity: 99.818

Percent Identity: 99.635

Length: 548

Gaps: 0

Percent Identity: 99.635

alignment_block:

US-09-581-241-6 x AR098258

Align seg 1/1 to: AR098258 from: 1 to: 1644

```

1 MetGlusMetGluAsnAspGluAsnIleValTYrGlyProGluProph 17
1 ATGGAAACATGGAGAACATGAAATATGTGTATGTCCTGAACCATP 50
17 eTYrProIleGluGluSerAlaGluIleAlaGluIleArgLysTYrMetA 34
51 TTACCTTATTGAGAGGAGATCTGCTGAGACACATTCGCGAATATATG 100
34 sPArgTYrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
101 ATCGATATGCCAAACTTGGAGCAATTCCTTACTACACGACTTACCGGT 150
51 ValAspTYrThrTYrAlaGluTYrIleGluLysSerCysCysLeuGly 67
151 GTGCGATTATACGTACGCCGAAATACCTTAGAAAAATCATGCTGTCTAG 200
67 uAlaLeuLysAsnTYrGlyLeuValValAspGlyArgIleAlaLeuCys 84
201 GCGTTTAAAGAAATTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 250
84 eArgLysAsnCysGluGluPhePheIleProValIleuAlaGlyLeuPhe 100
251 GTGAAACACTGTGAAGAAATCTTATCTCTGTATAGCCGTTTATTTATA 300
101 GlyValGlyValAlaProThrAsnGluIleTYrThrLeuArgGluLeuVa 117
301 GGTGTCGCTGCTGCTCACAATAAGATTAAGATTAACCTACGTAATGCT 350
117 IHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysG 134
351 TCACAGTTTAGCATCTCAACCAACAATTTGTAATTTACTTTCTAAAAAG 400
134 LysLeuAspLysValIleThrValGluLysThrValThrAlaIleLysThr 150
401 GATTAGATAAGTTATTAAGTACCAAAAAACCGTAACCTCTATTTAAACC 450
151 IleValIleLeuAspSerLysValAspTYrArgGlyTYrGlnSerMetAs 167
451 ATTGTTATATTGGACAGCAAAAGTGAATTAAGGTTATCAATTCATGGA 500
167 PAsnPheIleLysLysAsnThrProGluGlyPheLysGlySerSerPhe 184

```

```

501 CAACCTTATTAAAAAACAACCTCCACAGGTTTAAAGATCAAGTTTAA 550
184 ystrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
551 AAACGTAGAAAGTTAACCCGCAAGAACAAAGTGTCTTATATAGAACTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
601 TCGGGTTCACACCGGTTTGGCAAAAAGGTGTCACACTTACTATGAAATTT 650
217 eValThrArgPheSerHisAlaArgAspProIleTYrGlyAsnGlnValS 234
651 GGAGATACAGATTTTCTACCGCTAGAGATCCCAATTTATGGAACCAAGTT 700
234 eProGlyThrAlaIleLeuThrValValProPheHisGlyPheGly 250
701 CACACGACACGCGCTATTTTAACTGTAGTACCAATCCATCATGTTTGGT 750
251 MetPheThrThrLeuGlyTYrLeuThrCysGlyPheArgIleValMetLe 267
751 ATGTTTACTACTTTAGGCTATCTACTGTGTGTTTCTGATTCATGTT 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTYrLysC 284
801 AACGAAATTTGACGAGAGACCTTTTAAAAACACGCAACGATTTACAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAAGCGTTATCTTGAACCACTTGTGTCATTTCTTAATTAAGT 900
301 GluLeuLeuAspLysTYrAspLysSerAsnLeuValGluIleAlaSerG 317
901 GAATTTCTGATTAATATATGATTTATCAAAATTTGTTGAATTCATCTGG 950
317 xGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
951 CGAGACACCTTATCTTAAAGAAATGTGTGAAGCTGTGTGACGCTTTTA 1000
334 snLeuProGlyValArgGlnGlyTYrGlyLeuThrGluThrProSerAl 350
1001 ATTACCGGGGCTGCTCTCAAGCCTATGTTTAAACAAACACCTCTGCA 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
1051 ATTATTATCAACCGGAGCGGATGATTAACAGGTGCTTGGCAAAAT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThr 384
1101 TGTGCCATTATTTAAAGCAAAAGTTATGATCTGTGATCTTAAAAAATCT 1150
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
1151 TGGGCCCGAAGACAGCTGAGAAAGTTTGTGTAAGGCTCTATGCTTAAG 1200
401 LysGlyTYrValAspAsnProGluAlaThrArgGluIleIleAspGlu 417
1201 AAAGGTTATGTAGATTAATCCAGAACGACAAAGAAATCATGATGATAGA 1250
417 uGlyTYrPheLysHisThrGlyAspIleGlyTYrTYrAspGluGluLysHis 434
1251 AGGTGTTGTCACACGAGGATATTGGGTATTACGATGAAGAAAAACATT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTYrLysGlyTYrGln 450
1301 TCTTATAGTGTGATCGTTGAAGTCTTTATATCAATACAAAGATATCA 1350
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
1351 GTACACACCTGCTGAATTAGAATCTGCTTTTGAACAATCCAAATATTTT 1400
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484

```


1401 TGATGCCGCGCTGCTGCTGCTGCATCCTATAGCTGCTGAGCTCCGG 1450
484 1yalaValaValaValleuLysGlySerMetThrGluLysGluVal 500
1451 GAGCTGTTGTGTACTTAAAGAAAGAAATCTATGACTGAAAAAGAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgLysG 517
1501 ATGATTACAGTGTGCTAGTCAGATTCAAAATGCAAAACGTTGGCTGTGG 1550
517 ValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
1551 TGTCGCTTTTGGACGAGTAAAGTCTCAGCTGTAATAATGTGACG 1600
534 1yLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
1601 GTAAAGCAATTAGAGAAATACTGAAAGAACCGCTGCTAAGATG 1644
seq_name: gb_Pat:E02495
seq_documentation_block: 1644 bp RNA PAT 29-SEP-1997
LOCUS E02495 CDNA encoding Luciferase.
DEFINITION E02495
ACCESSION E02495.1 GI:2170725
VERSION JP 1990171189-A/1.
KEYWORDS
SOURCE Luciola lateralis.
ORGANISM Luciola lateralis.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
Blattelliformia; Cantharoidae; Lampyridae; Luciola.
REFERENCE 1 (bases 1 to 1644)
AUTHORS Kajiyama, N., Tatsumi, H. and Nakano, E.
TITLE LUCIFERASE GENE
PATENT: JP 1990171189-A 1 02-JUL-1990;
JOURNAL KIKKOMAN CORP
COMMENT
OS Luciola lateralis
PN JP 1990171189-A/1
PD 02-JUL-1990
PF 22-DEC-1988 JP 1988322029
PI KAJIYAMA NAOKI, TATSUMI HIROKI, NAKANO EICHI PC
C12N15/53/C12N9/02;
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH key
FT mat_peptide 1..1644
FT /product="Luciferase".
FEATURES
source Location/Qualifiers
1..1644
/organism="Luciola lateralis"
/db_xref="taxon:7052"
BASE COUNT 529 a 262 c 349 g 504 t
ORIGIN
alignment_scores:
Quality: 2814.00 Length: 548
Ratio: 5.144 Gaps: 0
Percent Similarity: 99.818 Percent Identity: 99.635
alignment_block:
US-09-581-241-6 x E02495 ..
Align seg 1/1 to: E02495 from: 1 to: 1644
1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
1 ATGGAAGAACATGAGAAACATGAAATATTTGTATAGTCTGACCAACCAT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMet 34

117 1yLysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
1601 GTAAAGCAATTAGAGAAATACTGAAAGAACCGCTGCTAAGATG 1644
seq_name: gb_Pat:E02495
seq_documentation_block: 1644 bp RNA PAT 29-SEP-1997
LOCUS E02495 CDNA encoding Luciferase.
DEFINITION E02495
ACCESSION E02495.1 GI:2170725
VERSION JP 1990171189-A/1.
KEYWORDS
SOURCE Luciola lateralis.
ORGANISM Luciola lateralis.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
Blattelliformia; Cantharoidae; Lampyridae; Luciola.
REFERENCE 1 (bases 1 to 1644)
AUTHORS Kajiyama, N., Tatsumi, H. and Nakano, E.
TITLE LUCIFERASE GENE
PATENT: JP 1990171189-A 1 02-JUL-1990;
JOURNAL KIKKOMAN CORP
COMMENT
OS Luciola lateralis
PN JP 1990171189-A/1
PD 02-JUL-1990
PF 22-DEC-1988 JP 1988322029
PI KAJIYAMA NAOKI, TATSUMI HIROKI, NAKANO EICHI PC
C12N15/53/C12N9/02;
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH key
FT mat_peptide 1..1644
FT /product="Luciferase".
FEATURES
source Location/Qualifiers
1..1644
/organism="Luciola lateralis"
/db_xref="taxon:7052"
BASE COUNT 529 a 262 c 349 g 504 t
ORIGIN
alignment_scores:
Quality: 2814.00 Length: 548
Ratio: 5.144 Gaps: 0
Percent Similarity: 99.818 Percent Identity: 99.635
alignment_block:
US-09-581-241-6 x E02495 ..
Align seg 1/1 to: E02495 from: 1 to: 1644
1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
1 ATGGAAGAACATGAGAAACATGAAATATTTGTATAGTCTGACCAACCAT 50
17 eTyrProIleGluGluGlySerAlaGlyAlaGlnLeuArgLysTyrMet 34

```

951 CGGAGCACCTTTATCTAAGAAATGGTGAAGCTTGTCTAGACGTTTTA 1000
334 snleupProglValArgInglYTrcIyleuThrgluThrThSerAla 350
|||||
1001 ATTTCACGGGTGTTGCGCAAGGCTATGTTTACAGAAAACACTCTGCA 1050
351 llelleleThrProgluIyAspAspLysProglYAlaSerGIyLysVa 367
|||||
1051 ATTATTATCACACCGGAAGCGCATGATAAACCAAGGTCCTCTGGCAAGT 1100
367 lValProleuPheLysAlaLysValIleAspLeuAspThrLysIsthrl 384
|||||
1101 TGTGCCATTTATTTAAAGCAAAAGTTATCGATCTTGATCTAAAAAACTT 1150
384 euGIyProAsnArgArgIyGIyValIcYsValIySGIyPrometLeuMet 400
|||||
1151 TGGGCCCGCAACACACGCGAGGAAGTTTGTCTAAAGGTCCTATGCTTATG 1200
401 LysGIyTYyValAspAsnProgluAlaThrArgIuIlelleAspGIuGI 417
|||||
1201 AAAGGTTATGTAGATATATCCAGAACACAAAGAAATCATAGATGAGA 1250
417 uGIyTrpLeuHisThrGIyAspIleGIyTYyTrpAspGIuGIuLysHisP 434
|||||
1251 AGGTTGTTGTCACACAGAGATATTTGGGATTTACGATGAAAGAAAACATT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTYyLysGIyTYyGI 450
|||||
1301 TCTTTATCGTGATCGTTGTAAGCTTTTATCAATAACAAAGATATNCA 1350
451 ValProProAlaGIuLeuGIuSerValLeuLeuGIuHisProAsnIlePh 467
|||||
1351 GTACCCACGCGTGAATTAAGATCTGTTTTCACATCCAAATATATTTT 1400
467 eaSpAlaGIyValAlaGIyValProAspProIleAlaGIyGIuLeuProG 484
|||||
1401 TGATGCCGGCGTTGCGGCTTCACATCTATAGCTGGAGCTTCCGG 1450
484 lYAlaValAlaValLeuLysLysGIyLysSerMetThrgIuLysGIuVal 500
|||||
1451 GAGCTGTTGTTGTAAGTGAAGAAAAGAAATCTATGACTGAAAAGAGA 1500
501 MetAspTYyValAlaSerGIuValSerAsnAlaLysArgLeuArgIyGI 517
|||||
1501 ATGGATTACGTTGCTAGTCAGATTCATAATGCAAAACGTTTGGCTGTG 1550
517 yValArgPheValAspGIuValProLysGIyLeuThrgIyLysIleAspG 534
|||||
1551 TGTCCGTTTGTGGAGAGAGTACTTAAGGTCTGCACTGTAATAATGACG 1600
534 lYlYsAlaIleArgGIuIleLeuLysLysProValAlaLysMet 548
|||||
1601 GTAAAGCAATTAAGAGAAATATGAAACCAAGTGTCTAAGATG 1644

```

seq_name: gb_pat:E05448

seq_documentation_block:

LOCUS	E05448	RNA	PAT	29-SEP-1997
DEFINITION	cDNA sequence of Luciferase.			
ACCESSION	E05448			
VERSION	E05448.1 GI:2173637			
KEYWORDS	JP 1993244942-A/2.			
SOURCE	Luciola lateralis.			
ORGANISM	Luciola lateralis			
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;			
	Elateriforma; Cantharoidea; Lampyridae; Luciola.			
REFERENCE	1 (bases 1 to 1644)			
AUTHORS	Kajiyama, N. and Nakano, E.			
TITLE	HEAT-STABLE FIREFLY LUCIFERASE, HEAT-STABLE FIREFLY LUCIFERASE			
JOURNAL	GENE, NEW RECOMBINANT DNA AND PRODUCTION OF HEAT-STABLE FIREFLY LUCIFERASE			
	Patent: JP 1993244942-A 2 24-SEP-1993;			

```

COMMENT
KIRKMAN CORP
OS Luciola lateralis
PN JP 1993244942-A/2
PD 24-SEP-1993
PE 22-MAY-1992 JP 1992131057
PR 27-JUN-1991 JP 91P 157117, 29-NOV-1991 JP 91P 317064 PI
PC C12N9/02.C12N15/53//C12N1/21.(C12N9/02.C12R1.19); CC
strandness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key location/Qualifiers
FT CDS 1..1644 /product='Luciola lateralis luciferase'.
FEATURES
source location/Qualifiers
1..1644 /organism='Luciola lateralis'
BASE COUNT 529 a 262 c 349 g 504 t
ORIGIN
alignment_scores:
Quality: 2814.00 Length: 548
Ratio: 5.144 Gaps: 0
Percent Similarity: 99.818 Percent Identity: 99.635
alignment_block:
US-09-581-241-6 x E05448 ..
Align seg 1/1 to: E05448 from: 1 to: 1644
1 MetGIuAsnMetGIuAsnAspGIuAsnIleValTYyGIyProGIuProPh 17
|||||
1 ATGGAAGAACATGGAGAACGATGAAATATGTTGTTATGCTCTGAACCAATT 50
17 eTYrProIleGIuGIySerAlaGIyValAGluLeuArgIyTYyMetA 34
|||||
51 TTTACCTATTGGAAGAGGATCTCTGAGCGACATTTGCCGACGTATATGG 100
34 sPArgTYrAlaLysIleuGIyAlaIleAlaPheThrAsnAlaLeuThrgIy 50
|||||
101 ATCGATATGCAAAACCTTGAGCATTTGCTTTTACTTAACGACACTTACCG 150
51 ValAspTYyThrTYrAlaGIuTYyLysGIuLysSerGIyLysIleuGIyGI 67
|||||
151 GTGGATTATACGACGCCGGAATATCTTAGAAAATCATGCTGTCTAGAGA 200
67 uAlaLeuLysAsnTYyGIyLeuValValAspGIyArgIleAlaLeuCys 84
|||||
201 GGCTTTAAAGAAATATGTTGGTGTGTGATGGAAGAAATTCGCTTATGCA 250
84 eGIuAsnGIySGIuGIuPhePheIleProValLeuAlaGIyLeuPheIle 100
|||||
251 GTGAAACGTGAGAGAAATCTTATTCCTGTATTTACCGGTTATTATTA 300
101 GIyValGIyValAlaProThrsnGIuIleTYyThrThrLeuArgIuLeu 117
|||||
301 GGTGTGCTGTGGCTCCAACTATGATTTACACTCTACGTAGGAATTTG 350
117 lHisSerLeuGIyIleSerLysProThrlleValPheSerSerLysGI 134
|||||
351 TCACAGTTTAGGCATCTCTAACCAACATTTGATTTACTTCTAAAGAAG 400
134 lYLeuAspLysValIleThrValGIuLysThrValThrAlaIleLysThr 150
|||||
401 GATTAGATTAAGTTATTAAGTACAAAAAACGTAACCTCTATTAAACC 450
151 lleValIleLeuAspSerLysValAspTYyArgGIyTYyGIuSerMetAs 167
|||||
451 ATTGTTATATTGGACAGCAAGTGAATTTAGAGTTATCAATCATCATGA 500

```

167 pAspNheIIeLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
 184 ysthrValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
 551 AAAGTGAAGAGTTAAACGCAAGAACAAAGTCTCTTAAATGAAGCTCT 600
 201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnI 217
 601 TCGGGTTCACACGGTTTGCACAAAGGTCGCACTTACATGCAAAATGC 650
 217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnVal 234
 651 AGTCACAGATGTTTCTCCAGCTAGATCCAAATTTATGAGAACCAAGTTT 700
 234 erProGlyThrAlaIleLeuThrValAlaProPheHisHisGlyPheGly 250
 701 CACCAAGCAGCGCTATTATTAACTGATGATCCATTCATGGTTTGGT 750
 251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIe 267
 751 ATGTTTACTACTTTAGGCTATCTAACTTGCGTTTCTGATATGTCATGTT 800
 267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
 801 AACCAAAATTTGACAGAGACTTTTAAAAACCTGCAAGATTAACAAT 850
 284 ysserSerValIleLeuValProThrPhePheAlaIleLeuAsnAsrSer 300
 851 GTTACAGCGTTATTCTTGTACCGACTTGTGTAATCTTAAATAGAAAGT 900
 301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
 901 GAATTAATCGATTAATATGATTATCAAAATTTAGTGAATTCATGATCGT 950
 317 yGlyAlaProLeuSerLysGluIleGlyAlaValAlaArgArgPheA 334
 951 CGAGCACCTTTATCTAAAGAAATTTGGTGAAGCTGTTGCTAGACCTTTA 1000
 334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
 1001 ATTACCGGGTTCGTCAGCATATGTTTAAACAGAAACAACCTGCA 1050
 351 IleIleIleThrProGlnGlyAspAspLysProGlyAlaSerGlyLysVa 367
 1051 ATTATTATCACACCGAAGCGATGATAAACAGGTGCTTCTGCAAGAT 1100
 367 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
 1101 TGTGCAATTAATTAAGCAAAAGTTATGATCTTATCTAAAAAACTT 1150
 384 euGlyProAsnArgArgLysGluValLysValLysGlyProMetLeuMet 400
 1151 TGGGCGCCGACAGACGTGAGAAAGTTGTGAAGAGGTCTATGCTTATG 1200
 401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGlnG 417
 1201 AAAGGTATTGTAGATATCCAGAACCAAGAACAAATCATAGATGAGAGA 1250
 417 uGlyTyrLeuHisThrGlyAspIleGlyTyrTyrAspGlnGluLysHisP 434
 1251 AGGTGCTGTTGCACACGAGATATTTGGGATTTACATGAGAGAAAAACTT 1300
 434 hspHelleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
 1301 TCTTTATCGTGCATGCTTTGAAGCTTTTAAATCAAAATACAAAGATATCAA 1350
 451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePh 467
 1351 GTACCAACGCTGTAATTAAGATCTGTTCTTTGACACATCCAAATATTTT 1400

467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyLeuLeuProG 484
 1401 TGATGCGGCGGTGCTGGCGCTCCAGATCTTACTTGTGACGTTCCGG 1450
 484 ValAlaValAlaValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
 1451 GAGCTGTTGTTGCTGCTGAAAAAGAAATATATGATGAGTAAAAAGACTA 1500
 501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyC 517
 1501 ATGGATTACGTTGCTGCTGCTCAAGTTTCAAAATGCAAAAGCTTGGTGG 1550
 517 yValArgPheValAspGlnValProLysGlyLeuThrGlyLysIleAspG 534
 1551 TGTGCTGTTTGTGGAGAGAGTACTTAAGTCTCACTGCTGTAATTTGACG 1600
 534 LysAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
 1601 GTAAAGCAATTAGAGAAATATGAGAAACAGTTGCTGAAGATG 1644

seq_name: gb_in:LLUCI

seq_documentation_block:

LOCUS LLUCI 1781 bp mRNA INV 24-NOV-1993
 DEFINITION L. lateralis luciferase mRNA.
 ACCESSION X66919.38389
 VERSION X66919.1 GI:9526
 KEYWORDS luciferase.
 SOURCE Japanese firefly.
 ORGANISM Luciola lateralis
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 Elateriformia; Cantharoidea; Lampyridae; Luciola.

REFERENCE
 1 (bases 1 to 1781)
 Tatsumi, H., Kajiyama, N. and Nakano, E.
 Molecular cloning and expression in *Escherichia coli* of a cDNA
 clone encoding luciferase of a firefly, *Luciola lateralis*

JOURNAL
 MEDLINE
 92305054
 Biochim. Biophys. Acta 1131 (2), 161-165 (1992)

REFERENCE
 2 (bases 1 to 1781)
 Tatsumi, H.
 Direct Submision
 Submitted (21-JUL-1992) H. Tatsumi, Research and Development
 Division, Kinokuni Corporation, 399 Noda, Noda City, Chiba 278,
 JAPAN

FEATURES
 source Location/Qualifiers
 1..1781

CDS

/organism="Luciola lateralis"
 /db_xref="taxon:7052"
 59..1705
 /codon_start=1
 /product="luciferase"
 /protein_id="CAA47358.1"
 /db_xref="GI:9527"
 /db_xref="SWISS-PROT:O01158"
 /translation="MEMENDENTVYGEPEPTPIEBESAGQLKKYDRYAKLCAINP
 TNALGVDTYTAELVLEKSCLEALKNGLVVDGIALVCSNCEEPFVPLAGLFGV
 GVAPEINEYTLLELVHLSGSKPTLVESKGLDKVITVQKTVIAKIVYLLDSKVY
 RGVQMDNFKKTPQGRKSGSEFTEVNRKEQVALINSSGSLGPGVOLTDENAV
 TRFSHARDPIYCNQVSPGTAITLVVPHFGGMFTLLGLTCGPRIVMLTKFDETEL
 KTLQDYKCSVLTLPYTLALNRSLLDKYLSNVEIASGCAPLSKETGAVARRFN
 LRGVNRGGLVETTSAILITTEPGDUPKASGVVPLFRAKVIIDDTKTKLGPNRGEV
 CVKGPMLKMGVNDPEATREIIDEQGLHDTDIDGVEDEKHFIVDRKLSLKRTGTO
 VPAELESVLLQHPNIFDAGVAGVDPPIAGELPGAVVLEKGSMTKEVDVYASOV
 SNAKRLRGVREVDEVPGLTGKIDGKAIKREILKKPVAKM"

BASE COUNT

579 a 276 c 367 g 559 t

alignment_scores:

Quality: 2814.00 Length: 548
 Ratio: 5.144 Gaps: 0
 Percent Similarity: 99.818 Percent Identity: 99.635

alignment block:

US-09-581-241-6 x LLLUCI ..

Align seg 1/1 to: LLLUCI from: 1 to: 1781

```

1 MetGluAsnMetGluAsnAspGluAsnIleValuYrGlyProGluProph 17
  |||||||
59 ATGGAAACACTGGAGAACGATGAAATATGCTATGCTGCAACCATTT 108
  |||||||
17 eYrProIleGluGlySerIleGluAlaGluIleuArgLysTyrMet 34
  |||||||
109 TTAACCTATGAGAGAGGATCTCTGGAGCACAATTGCCAATATATGG 138
  |||||||
34 sparGyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
  |||||||
159 ATCGATATGCAAAACTTGGAGCAATGCTTTTACTAACGCACCTTACCGGT 208
  |||||||
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysGlyLeuGly 67
  |||||||
209 GTGATATATACGTCACCCGATATCTAGAAAATCATGCTGTCTAGAGGA 258
  |||||||
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
  |||||||
259 GCGTTTAAAGAAATTATGTTGGTGTGTGATGAGAGAATTCGCTATGCA 308
  |||||||
84 eGluAsnGlyGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
  |||||||
309 GTGAAACGTCGAGAAATCTTATTCCTGTATTCAGCCGCTTATTTATA 358
  |||||||
101 GluValGlyValAlaProThrAsnGluIleYrThrLeuArgGluLeuVal 117
  |||||||
359 GGTGTCGGTGGTCCACATATGATGATTTACACTACGTGATGATTTGGT 408
  |||||||
117 LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLysG 134
  |||||||
409 TCACAGTTAGGCATCTCTAACCCAACTGATTTAGCTCTCAAAAAAG 458
  |||||||
134 LLeuAspLysValIlePheValGluLysThrValThrAlaIleLysThr 150
  |||||||
459 GATTTGATTAAGTTATTAAGTACCAAAAAACGTAACCTCTTTAAAAAC 508
  |||||||
151 LLeuAlaIleLeuAspSerLysValAspTyrArgIleYrGlnSerMet 167
  |||||||
509 ATTGTTAATTTGACAGCAAGTGTATAGAGTTATCAATCCATGCA 558
  |||||||
167 PAsnPheIleLysAsnThrProGlnGlyPheLysGlySerSerPhe 184
  |||||||
559 CAACCTTTATTAATAAAACACTCCACAAGGTTTCAAAAGATCAAGTTTAA 608
  |||||||
184 YSThrValGluValAsnArgLysGluGluValAlaLeuIleMetAsnSer 200
  |||||||
609 AAACGTGAGAGTTAACCCGAAAGAAAGTTCGCTTATATATGAACTCT 658
  |||||||
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsn 217
  |||||||
659 TCGGGTTACCGGTTTCCAAAAGGTGTGCAACTTACTCATGAAAGAAC 708
  |||||||
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGluVal 234
  |||||||
709 AGTCACATGATTTTCTCACGCTAGATCCAAATTTATGAAACCAAGTTT 758
  |||||||
234 ePProGlyThrAlaIleLeuThrValValIProPheHisHisGlyPheGly 250
  |||||||
759 CACACGACGCGCTATTTTAACTGATAGACATTCATCCATCATGCTTTGGT 808
  |||||||
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
  |||||||
809 ATGTTTACTACTTATGCTATCTAATCTTGTGTTTCTGATTTGATGATTT 858
  |||||||
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
  |||||||
859 AAGCAATTTGACGAGAGACTTTTAAAAACACTGCAAGATTTACAAAT 908
  |||||||

```

```

284 YSerserValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
  |||||||
909 GTTCAACCGTATTTCTGTACCACTTTGTTGCAATTTCTTAATAGAGT 958
  |||||||
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
  |||||||
959 GAATTTACTGATTAATATGATTTATCAAAATTTAGTGAATTCATCTCG 1008
  |||||||
317 YGlyAlaProLeuSerLysGluIleGlyAlaValAlaArgArgPhe 334
  |||||||
1009 CGGAGCACCTTTATCTAAAGAAATGGTGAAGCTGTGTCAGACGTTTAA 1058
  |||||||
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
  |||||||
1059 ATTACCGGGTGTTCGCAAGCTATGTTTAAACAAACACCTCTGCA 1108
  |||||||
351 LLeIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
  |||||||
1109 ATTATTTATCACACCGGAGCGATGATMAACCAAGCTCTTCTGCAAGT 1158
  |||||||
367 LValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
  |||||||
1159 TGTCCCATTTATTTAAAGCAAAAGTTATCGATCTTGAATAAAAAACTT 1208
  |||||||
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
  |||||||
1209 TGGGCCCGCAACAGACGTGAGAAAGTTTGTGTAAGGCTCCTATGCTTANG 1258
  |||||||
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleLeuAspGlu 417
  |||||||
1259 AAGGTTATATGATATATCCAGAGCAACAGAGAAATATATGATGATAA 1308
  |||||||
417 uGlyTyrLeuHisThrGlyAspIleGlyTyrTyrAspGluGluYrHisP 434
  |||||||
1309 AGGTTGTTGCAACAGAGATATGGTATTAACGATGAAGAAAAACATT 1358
  |||||||
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
  |||||||
1359 TCTTATATCGATCGTTGGAAGCTTTTAAATCAAAATACAAAGATATCAA 1408
  |||||||
451 ValProProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIleP 467
  |||||||
1409 GTACACACCGTCGATTAATCAATCTGTTTTCGCAACATCCAAATTTT 1458
  |||||||
467 eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
  |||||||
1459 TGATGCGCGCTTGTGCGCTCCAGATCTTATAGCTGTGAGCTCCGG 1508
  |||||||
484 LValValValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
  |||||||
1509 GAGCTGTTGTTTACTTGAAAGAAAGAAATCTATGACTGAAAAAAGACTA 1558
  |||||||
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
  |||||||
1559 ATGATATAGCTGTGATGCAAGTTTCAAAATGCAAAACGTTCCGAGTGG 1608
  |||||||
517 YValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
  |||||||
1609 TGTCCGTTTGTGCGAGCAAGTACCTTAAAGGTCCTCACTGTAATAATGACG 1658
  |||||||
534 LLeuAlaIleArgGluIleLeuLysLysProValAlaLysMet 548
  |||||||
1659 GTAAAGCAATTTAGAGAAATACTGAAAGAAACGAGTTGCTAAGATG 1702
  |||||||

```

seq_name: gb_pat:AR043317

seq_documentation_block:

LOCUS AR043317 1908 bp DNA PAT 29-SEP-1999

DEFINITION Sequence 8 from patent US 5614465.

ACCESSION AR043317

VERSION AR043317.1 GI:5964325

KEYWORDS


```

501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgLysG1 517
|||||
1501 ATGATTACCTTCTGTAGTCAAGTTTCAAAAGTTCGAACAAAGTTGGGTGG 1550
517 yValArgPheValAspGlnValProLysGlyLeuThrGlyLysIleAspG 534
|||||
1551 TGTCCGTTTGTGGACGAAGTACCTAAAGGCTCAGCTGTAATAATTGACG 1600
534 yLysAlaIleArgGlnIleLeuLysLysProValAlaLys 547
|||||
1601 GTAAGCAATTAAGAAATTAATCTAAGAAACCACTGCTGAAG 1641

seq_name: gb_pat:AR062709

seq_documentation_block:
LOCUS AR062709 1908 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 8 from patent US 5843746.
ACCESSION AR062709
VERSION AR062709.1 GI:5990400
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
REFERENCE
1 (bases 1 to 1908)
Tatsumi,H., Fukuda,S., Kikuchi,M. and Koyama,Y.
Biotinlated firefly luciferase, a gene for biotinlated firefly
luciferase, a recombinant DNA, a process for producing biotinlated
luciferase and a bioluminescent analysis method
Patent: US 5843746-A 8 01-DEC-1998;
FEATURES
LOCATION/Qualifiers
SOURCE 1..1908
BASE COUNT 598 a 329 c 426 g 555 t
ORIGIN

alignment_scores:
Quality: 2812.00 Length: 547
Ratio: 5.141 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.634

alignment_block:
US-09-581-241-6 x AR062709 ..
Align seg 1/1 to: AR062709 from: 1 to: 1908

```

```

117 lHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLysG 134
|||||
351 TCACAGTTTAGGCATCTCTAAGCCAAACAAATTTATTTAGTTCTTAAGAAAG 400
134 yLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
|||||
401 GATTACATAAAGTTATTAAGTCAACAAACGTAAGTCTGCTATTAAGCAACC 450
151 lIleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
|||||
451 ATTGTTATATTATGGACAGCAAGATGATATAGAGTTATCAATCAATCAGCA 500
167 PAspPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheL 184
|||||
501 CAACCTTATTAATAAATAACACCTCACAGAGTTTCAAAAGATCAAGTTTAA 550
184 yThrValGlnValAsnArgLysGlnValAlaIleLeuIleMetAsnSer 200
|||||
551 AAACGTAGAGTTAACCGCAAGAACAAGTCTCTTTAATGAACCTCT 600
201 SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGlnAsnI 217
|||||
601 TCGGTTCAACCGGTTTGCCAAAGGTGTCACACTTACATGAATAATTT 650
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
|||||
651 GGTACACCGCTTTTCTCAGCTAGAGATCCAAATTTAGAAACCAAGTTT 700
234 eProGlyThrAlaIleLeuThrValAlaProPheHisHisGlyPheGly 250
|||||
701 CACACAGCACGGCTATTTTAACTGATGACATTCACATCATGATGTTTGGT 750
251 MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIe 267
|||||
751 ATGTTTACTACTTAAAGCTATCTAATCTGATGTTTCCATTTGTCATGTT 800
267 uThrLysPheAspGlnGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
|||||
801 AACGAAATTTTGACGAAGAGACCTTTTAAACACACGCAAGATTACAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
|||||
851 GTTCAAGCGTTATCTTGTACCGACTTGTGTCAAATCTTAATAGAAAT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGlnIleAlaSerG1 317
|||||
901 GAATTAAGTGAATAATATGATTTATCAAAATTTAGTTGAATATGCAATCTGG 950
317 yGlyAlaProLeuSerLysGlnIleGlyGlnAlaValAlaArgArgPheA 334
|||||
951 CGAGCACCTTATCTAAAGAAATGAGTGAACCTGTTGCTAGACGTTTAA 1000
334 snLeuProGlyValAlaArgGlnGlyTyrGlyLeuThrGluThrHisSerAla 350
|||||
1001 ATTTACCGGCTGTTGTCAAAGCTATGTTTAAACAGAAACCAACCTTCGA 1050
351 lIleIleIleThrProGlnGlyAspAspLysProGlyAlaSerGlyLysVa 367
|||||
1051 ATTATTATCACACCGCAAGCGCATATTAACCAAGTCTCTTGGCAAAAGT 1100
367 lValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
|||||
1101 TGTGCATTTATTTAAAGCAAAAGTTATGATGATCTGATCAAAAAAACTT 1150
384 eGlyProAsnArgArgGlyGlnValAlaLysValLysGlyProMetLeuMet 400
|||||
1151 TGGGCCCAAGACAGCTGGAGAAAGTTGTGTAAGGCTCATGCTTATG 1200
401 LysGlyTyrValAspAsnProGlnuAlaThrArgGlnIleIleAspGlnG1 417
|||||
1201 AAAGTTATGTATGATTAATCCAGAGCAACAGCAAAATCATATGATGAAGA 1250

```

```

417 uGlyTrpLeuHisThrGlyAspIleGlyTyrTyrAspGluGlyHisP 434
|||||
1251 AGTTGGTGCACACAGAGATATGGTATTCAGATGAAGAAAAACATT 1300
|||||
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
|||||
1301 TCTTATTCGTCGATCGCTTGAAGCTTTATCAATACAAAGATATCA 1350
|||||
451 ValProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIleP 467
|||||
1351 GTACCACCTGCTGAATTAGATCTGCTTTTGCACATCCAAATATTTT 1400
|||||
467 eaSPAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuPro 484
|||||
1401 TGATGCCGGCGCTGCTGCGCTCCAGATCCATAGCTGGTGGCTTCGG 1450
|||||
484 IValAlaValAlaLeuLysGlyLysSerMetThrGluLysGluVal 500
|||||
1451 GAGCTGTTGTTGCTACTGAAAAAGAAAAATCTATGACTGAAAAAGAGTA 1500
|||||
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGly 517
|||||
1501 ATGATTAACGTTGCTAGTCAAGTTCAATGCAAAACGTTGCGTGGTGG 1550
|||||
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAsp 534
|||||
1551 TGTCGGTTTGGACGAGTACTAAAGCTCTCAGTGTGTAATTTAGC 1600
|||||
534 IYLSAlaIleArgGluIleLeuLysLysProValAlaLys 547
|||||
1601 GTAAAGCAATTAGAGAAATFACTGAAGAACCAAGTTGCTAAG 1641
|||||

seq_name: gb_pat:E12279
seq_documentation_block:
LOCUS E12279 1908 bp DNA PAT 24-JUN-1998
DEFINITION CDNA encoding biotinylated Luciola luciferase.
ACCESSION E12279
VERSION E12279.1 GI:3251113
KEYWORDS JP 1996308578-A/4.
SOURCE Luciola lateralis.
ORGANISM Luciola lateralis.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
Elateriformia; Cantharoidae; Lampyridae; Luciola.
REFERENCE
1 (bases 1 to 1908)
Tatsumi, H., Fukuda, M., Kikuchi, M. and Koyama, T.
BIOTIN FIRE FLY LUCIFERASE, BIOTIN FIRE FLY LUCIFERASE GENE, NEW
RECOMBINANT DNA, PRODUCTION OF BIOTIN FIRE FLY LUCIFERASE AND
BIOLUMINESCENCE ANALYSIS
Patent: JP 1996308578-A 4 26-NOV-1996:
JOURNAL KIKKOMAN CORP
COMMENT
OS Luciola lateralis
PN JP 1996308578-A/4
PD 26-NOV-1996
PR 24-APR-1995 JP 1995098857
PR 27-JUL-1994 JP 94P 193798, 14-MAR-1995 JP 95P 54625 PI
TATSUMI HIROKI, FUKUDA MASARU, KIKUCHI MANORU, KOYAMA TAIJI PC
C12N15/09,C12N9/02,C12O1/26,(C12N9/02,C12N1:185); CC
strandedness: Double;
CC topology: Unknown;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers
FT source 1. 1908
FT /clone='pHLf248' /organism='Luciola lateralis' FT
FT mat_peptide 1. 1908
FT /product='biotinylated Luciola luciferase'.
FEATURES
Location/Qualifiers
1. 1908
source /organism='Luciola lateralis'

```

```

BASE COUNT 598 a 329 c 426 g 555 t
ORIGIN
alignment_scores:
Quality: 2812.00 Length: 547
Ratio: 5.141 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.634
alignment_block:
US-09-581-241-6 x E12279
Align seg 1/1 to: E12279 from: 1 to: 1908
1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProPh 17
|||||
1 ATCGAAACATGAGCAAGCATGCAAAATATGGTATGATGCTGGAACATT 50
|||||
17 eYrProIleGluGluGlySerAlaGlyAlaIleAlaPheThrAsnAlaLeuThrGly 34
|||||
51 TTACCTATTGAAGAGGATCTCTGGACACACATTGCCCAAGTATATGG 100
|||||
34 sPaArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
|||||
101 ATCGATATGCAAAACTGGACCAATTGCTTTACTAACGCACATTACCGGT 150
|||||
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysLeuGly 67
|||||
151 GTGCAATTATACGTACGCCGCAATCTTAGAAAAATCAATGCTGTCTAGACA 200
|||||
67 uAlaLeuLysAsnTyrGlyLeuValAlaAspGlyArgIleAlaLeuCys 84
|||||
201 GCCTTTAAAGAAATATGCTTGTGTTGATGAGAAATTCGCTTATGCA 250
|||||
84 eRgLuAsnCysGluGluPhePheIleProValLeuAlaGlyLeuPheIle 100
|||||
251 GTGAAACCTGTGAAGAGTCTTATTCCTGATTTACCGGTTATTTATA 300
|||||
101 GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
|||||
301 GGTCGTGCTGTGCTCCACCTAAATGAGATTACACTACCTACGTAATTTGGT 350
|||||
117 LHisSerLeuGlyIleSerLysProThrIleValAlaPheSerSerLysGly 134
|||||
351 TCACAGTTTAGCATCTCTAACGCCAACAAATGTATTAGTTCAAAAAG 400
|||||
134 IYLeuAspLysValIleThrValGlnLysThrValThrAlaIleLysThr 150
|||||
401 GATTAGATAAAGTTATTAACGTACAAAAACGGTAACGTCTATTATAAAC 450
|||||
151 IleValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
|||||
451 ATTGTTATATTGGACACCAAGTCGATTATAGAGCTTATCAATCCATGGA 500
|||||
167 PaSnPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheIle 184
|||||
501 CAACTATTATTAATAAAAAACACTCCACAAGTTTCAAAAGATCAAGTTTAA 550
|||||
184 yStrValAlaGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
|||||
551 AAACGTAGAAAGTTAACCGCAAAAGAACAAAGTGTCTTATAAAGAACTCT 600
|||||
201 SerGlySerThrGlyLeuProLysGlyValAlaGlnLeuThrHisGluAsnIle 217
|||||
601 TCGGCTTCAACCGGTTTGCACAAAGGTGTGCACTTCTCATGATAAATTT 650
|||||
217 eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
|||||
651 GGTCAACGGGTTTTCACGCTAGACATCAATTTATGAAACCAAGTTT 700
|||||
234 eRrProGlyThrAlaIleLeuThrValAlaProPheHisHisGlyPheGly 250
|||||

```



```

701 CACAGGACGCGCTATTCTAGTACCATTCATCATGTTTGGT 750
251 MetPheThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetIle 267
751 ATGTTTACTACTTAGGCTATCTTAACCTTGCGTTTGTGATGTCAGTTP 800
267 uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
801 AACGAATTTGACGAGAGACTTTTAAAAACACCTGCACGATTTACAAAT 850
284 ySerSerValIleLeuValProThrLeuPheAlaIleLeuAsnArgSer 300
851 GTTCAAGCGTTATCTTGTAACGACTTGTTCATTTCTTAATAGAGT 900
301 GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleLeuAspG 317
901 GATTTCTGCTAATATATGATTTATCAAAATTTAGTTGAAATTCATCTGG 950
317 yGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgPhe 334
951 CGAGACGACTTATCTAAAGAAATGTGTAAGCTGTGCTAGACGTTTGA 1000
334 snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
1001 ATTTACCGGGTGTTCGTCAGAGGCTATGTTTAAACAGAAACACCTCTGCA 1050
351 IleIleIleThrProGluGlyAspAspLysProGlyAlaSerGlyLysVa 367
1051 ATTATATATACACCGGAGCGCATTAACCGAGTCTCTGCGCAAGT 1100
367 yValProLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrL 384
1101 TCGCCATTATTTAAAGCAAAAGTTTCATCTGTACTTAAAAAAAAGCTT 1150
384 euGlyProAsnArgArgGlyGluValCysValLysGlyProMetLeuMet 400
1151 TGGCCCGCAACAGAGCTGGAAGATTGTGTAAAGGCTCCTATGCTTATG 1200
401 LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluG 417
1201 AAGGTTAGTATGATTAATCCAGAACGACAGCAAGAAATCATGATGATGAA 1250
417 uGlyTrpLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisP 434
1251 AGGTTGGTTCACACAGAGATTTGGGTAATTACGATGAAGAAAAACATT 1300
434 hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
1301 TCTTATCGTGGATCGTTGGAAGTCTTTAATCAAAATACAAAGGATATCA 1350
451 ValProProAlaGluLeuGlnSerValLeuLeuGlnHisProAsnIlePh 467
1351 GTACACCGCTGGAATTAAGATCTGTTCTTTGCAACATCCAAATAATTTT 1400
467 eaSpAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuProG 484
1401 TGATCCCGGCGGTGCTGCGCTCCAGATCTTATAGCTGCTGAGCTTCCGG 1450
484 yAlaValValLeuLysLysGlyLysSerMetThrGluLysGluVal 500
1451 GAGCTGTGTGTTACTTGAAAAAGAAAAATCTATGACTGAAAAAGAGTA 1500
501 MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyG 517
1501 ATGGATTACGTTGCTAGTCAAGTTTCAAAATGCAAAACGTTTGCCTGGTGG 1550
517 yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
1551 TGTCGCTTTTGGAGCAGAGTACCTAAAGGCTCTACTGTTAAATTTGACG 1600
534 yLysAlaIleArgGluIleLeuLysLysProValAlaLys 547
1601 GTAAAGCAATTAGAGAAATACTGAGAAACCGAGTGTCTAAG 1641

```

```

seq_name: gb_pat:E16288
seq_documentation_block:
LOCUS E16288 1920 bp DNA PAT 28-JUL-1999
DEFINITION DNA encoding fusion protein which comprises luciferase and
lysostaphin.
ACCESSION E16288
VERSION E16288.1 GI:5710971
KEYWORDS JP 1998150991-A/1.
SOURCE JP 1998150991-A/1.
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 1920)
AUTHORS Tatsumi, H., Fukuda, M. and Nagahara, A.
TITLE LUCIFERASE-LYOSTAPHIN FUSED PROTEIN, ITS PRODUCTION AND
BIOLOGICAL ANALYSIS USING THE SAME
JOURNAL patent: JP 1998150991-A 1 09-JUN-1998;
KIKKOMAN CORP
COMMENT
OS None
OC Artificial sequences.
PN JP 1998150991-A/1
PD 09-JUN-1998
PF 25-NOV-1996 JP 1996328042
PI TATSUMI HIROKI, FUKUDA MASARU, NAGAHARA AYUMI PC
C12N15/09, C07H21/04, C07K19/00, C12N9/02, C12N9/52, C12P21/02, PC
C12Q1/66
PC (C12N9/02, C12R1:19), (C12N9/52, C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..1920 /organism="Artificial sequences".
FT location/Qualifiers
1..1920
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 627 a 310 c 407 g 576 t
ORIGIN
alignment_scores:
Quality: 2812.00 Length: 547
Ratio: 5.141 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.634
alignment_block:
US-09-581-241-6 x E16288
Align seg 1/1 to: E16288 from: 1 to: 1920
1 MetGluAsnMetGluAsnAspGluAsnIleValTyrGlyProGluProh 17
|||||
1 ATGGAACACATGAGAGACGATGATAATTTGCTATGCTGCAACCAT 50
17 eTyrProIleGluGluGlySerAlaGlyValGlnLeuArgLysTyrMet 34
|||||
51 TTACCTTATGGAAGAGGATCTGCTGAGCACAAATTTGCCAAGTATAG 100
34 spArgTyrAlaLysLeuGlyAlaIleAlaPheThrAsnAlaLeuThrGly 50
|||||
101 ATCGATTATGCAAACTTGAGCAATTTGCTTACTAGACGCACTTACCG 150
51 ValAspTyrThrTyrAlaGluTyrLeuGluLysSerCysGlyLeuGly 67
|||||
151 GTCGATTATACGTACGCCGAATATCTAGAAAAATCATGCTGCTAGAGA 200
67 uAlaLeuLysAsnTyrGlyLeuValValAspGlyArgIleAlaLeuCys 84
|||||
201 GCGTTTAAAGATTTATGTTGTTGTTGATGAGAAATTTGCTTATAGCA 250

```



```

84  erGluasnCySGluGluPhePheIleProValIleuAlaGlyLeuPheIle 100
    |||
251  GTGAAGCTGAGAGAGTCTTATTCCTGATATGACCGGTTATATTATA 300
    |||
101  GlyValGlyValAlaProThrAsnGluIleTyrThrLeuArgGluLeuVal 117
    |||
301  GGTGTGGTGGCTCCAACTAATGAGATTACCTGACGTGATGTGT 350
    |||
117  LHisSerLeuGlyIleSerLysProThrIleValPheSerSerLysLysG 134
    |||
351  TCACAGTTAGGACATCTCTAAGCCACAATGTTAGTTCTATAAAG 400
    |||
134  LLeuAspLysValIleThrValIleLysThrValIleAlaIleLysThr 150
    |||
401  GATTAGATTAAGTTATTAAGTACAAAAAGGTAACCTGCTATTAAAAAC 450
    |||
151  LLeuValIleLeuAspSerLysValAspTyrArgGlyTyrGlnSerMetAs 167
    |||
451  ATTGTTATATTGGACAGCAAGTGGATTATGAGGTTATCATCATGGA 500
    |||
167  PAspPheIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheI 184
    |||
501  CAACCTTTATTAATAAAAAACACTCCACAAGGTTTCAAGATCAAGTTT 550
    |||
184  ySThValGluValAsnArgLysGluGlnValAlaLeuIleMetAsnSer 200
    |||
551  AAACGTGTGAAGTTAACCGCAAGCAAGTTGCTTATTAATGAACCTCT 600
    |||
201  SerGlySerThrGlyLeuProLysGlyValGlnLeuThrHisGluAsnI 217
    |||
601  TCGGTTTACCGGTTTCCAAAGGTGTGCACTTACTCATGAATAATT 650
    |||
217  eValThrArgPheSerHisAlaArgAspProIleTyrGlyAsnGlnValS 234
    |||
651  GGTACAGCGTTTCTCAGCTAGAGATCCAAATTATGAAACCAAGTTT 700
    |||
234  eProGlyThrAlaIleLeuThrValIleProPheHisGlyPheGly 250
    |||
701  CACCGAGGACGGCTATTTTAACTGATGACATTCATCATGCTTTTGT 750
    |||
251  MetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLe 267
    |||
751  AGCTTTACACTTATGAGCTATCTACTGTTGTTGTTGTTGTTGTTGTT 800
    |||
267  uThrLysPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysC 284
    |||
801  AACGAAATTTGACGAAGAGACTTTTAAAAACACTGCAGATTAACAAT 850
    |||
284  ySerSerValIleLeuValIleProThrLeuPheAlaIleLeuAsnArgSer 300
    |||
851  GTTCAAGCGTTTATCTTCTGACGACTTGTGTCATTTCTTAATGAGAGT 900
    |||
301  GluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerG 317
    |||
901  GAATTACTCGATAAATATGATTTATCAAAATTTAGTTGAATTCATCTGG 950
    |||
317  yGluAlaProLeuSerLysGluIleGlyGluAlaValAlaArgAspPheA 334
    |||
951  CGAGACACCTTATCTAAAGAAATGTTGTAAGCTGTTGTAACGCTTTTA 1000
    |||
334  snLeuProGlyValArgGlnGlyTyrGlyLeuThrGluThrThrSerAla 350
    |||
1001  ATTTACCGGGTGTGCTCAAGCTATGCTTAAACAGAAACAACCTCTGCA 1050
    |||
351  IleIleIleThrProGluLysAspLysProGlyAlaSerGlyLysVal 367
    |||
1051  ATATTATATCACACCGGAGCGCATGATTAACACAGGTCCTTCGCAAACT 1100
    |||
367  ValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrL 384
    |||
1101  TGTGCCATTTATTAAGCAAAAGTTATCATCTCTGATTAATAAAAAACTT 1150

```

```

384  euGlyProAsnArgArgGlyValIleCysValLysGlyProMetLeuMet 400
    |||
1151  TGGGCCGAGACAGAGCTGGAGAAAGTTTGTCTAAAGGCTCTTACTTAA 1200
    |||
401  LysGlyTyrValAspAsnProGluAlaThrArgGluIleIleAspGluI 417
    |||
1201  AAGGTTATGTAGATATATCCAGAACGACACAGAAAGAAATCATGATGA 1250
    |||
417  uGlyThrPheHisThrGlyAspIleGlyTyrTyrAspGluGluLysHis 434
    |||
1251  AGGTTGTTGTCACACAGAGATATGGTATTAAGTGAAGAAAAACATT 1300
    |||
434  hePheIleValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGln 450
    |||
1301  TCTTTATCGGATCGCTTTGAAGCTCTTATATCAAAATACAAAGATATCA 1350
    |||
451  ValProProAlaGluLeuGluSerValIleLeuGlnHisProAsnIleP 467
    |||
1351  GTACCACTGCTGAATTAAGATCTGTTTGTGCAACATCCAAATATTTT 1400
    |||
467  eAspAlaGlyValAlaGlyValProAspProIleAlaGlyGluLeuPro 484
    |||
1401  TGAATGCCGGCTGCTGGCTCCAGATCTATAGCTGCTGACCTTCCGG 1450
    |||
484  LValAlaValIleValLysGlyLysSerMetThrGluGluVal 500
    |||
1451  GAGCTGTTGTGTACTTGAAAAAGAAATCTATGACTGAAAAAGAAAT 1500
    |||
501  MetAspTyrValAlaSerGlnValSerAsnAlaLysArgLeuArgLys 517
    |||
1501  ATGATTAACGTTGCTGATGTCAGATTCAAATGCAAAAGCTTGGTGGTGG 1550
    |||
517  yValArgPheValAspGluValProLysGlyLeuThrGlyLysIleAspG 534
    |||
1551  TGTCCGTTTGTGACGACAACTACTTAAAGGTCCTCAGCTGTAATTTGACG 1600
    |||
534  LysAlaIleArgGluIleLeuLysLysProValAlaLys 547
    |||
1601  GTAAAGCAATTAAGAAATACTGAAAGAACCAAGTGTCTAAG 1641
    |||
seq_name: gb_pat:AR043316

seq_documentation_block:
LOCUS AR043316 1704 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 5 from patent US 5814465.
ACCESSION AR043316
VERSION AR043316.1 GI:5964324
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1704)
AUTHORS
Tatsumi, H., Fukuda, S., Kikuchi, M. and Koyama, Y.
TITLE
Biotinlated firefly luciferase, a gene for biotinlated firefly
luciferase, a recombinant DNA, a process for producing biotinlated
luciferase and a bioluminescent analysis method
JOURNAL
Patent: US 5814465-A 5 29-SEP-1999;
FEATURES
location/Qualifiers
1..1704
BASE COUNT 541 a 275 c 361 g 527 t
ORIGIN
alignment_scores:
Quality: 2799.00 Length: 546
Ratio: 5.126 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.267

alignment_block:
US-09-581-241-6 x AR043316 ..
Align seg 1/1 to: AR043316 from: 1 to: 1704

```

[illegible]

TITLE Biocinated firefly luciferase, a gene for biocinated firefly

luciferase, a recombinant DNA, a process for producing biocinated

luciferase and a bioluminescent analysis method

JOURNAL Patent: US 5843746-A 5 01-DEC-1998;

FEATURES Location/Qualifiers

source 1..1704

BASE COUNT 541 a 275 c 361 g 527 t

ORIGIN /organism="unknown"

alignment_scores:

Quality: 2799.00 Length: 546
Ratio: 5.126 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.267

alignment_block:

US-09-581-241-6 x AR062708 ..

Align seg 1/1 to: AR062708 from: 1 to: 1704

```

3 AsnMetGlusnaspgluasnlllevaltyrGlyProGluProphetyr 19
:.....
67 AGTCTCGAAGACGATGAAATATTGTCTATGCTCGAACCATTTTACCC 116
19 oileGluGluGlySerAlaGlyAlaGluLeuArGlySerMetAspar 36
|||||
117 TATTGAAGAGGATCTCTGAGCACAATTGCCAATATATGATGAT 166
36 yralAlaLeuGlyAlaAlaAlaPheThrAsnAlaLeuThrGlyValAsp 52
|||||
167 ATGCAAAACTGGAGCAATTGCTTTACTAACGCACCTTACCGGTGCGAT 216
53 TyrThrTyraIaGluTyrluGluLysSerCysCysLeuGlyAlaLe 69
|||||
217 TATACGTACGCCGCAATCTTAGAAATATCAATGCTGAGAGGCTTT 266
69 uLysAsnTyrlGlyLeuValValAspGlyArgIleAlaLeuCysSerGlu 86
|||||
267 AAAGATATTATGTTGGTTGGTTGATGAGAGAAATGCGTTATGCGTAAA 316
86 snCysGluGluPhePheIleProValLeuAlaGlyLeuPheIleGlyAl 102
|||||
317 ACTGGAAGAGTCTTATCTGATTTAGCGGGTTATTATTATAGGGTGC 366
103 GlyAlaIaProThrAsnGluIleTyrThrLeuArgIleuValHisSe 119
|||||
367 GGTGGGCTCCAACTAATGAGATTACCTACGTGATTTGGTTCACAG 416
119 rleuGlyIleSerLysProThrIleValPheSerSerLysLysGlyLeu 136
|||||
417 TTTAGCATCTCTAAGCCACACATTTGATTAGTTCTAAAAAAGATTAG 466
136 sPLysValIleThrValGluLysThrValThrAlaIleLysThrIleVal 152
|||||
467 ATAAAGTTATACGTACAAAAAAGCGTAACTGCTATTTAAAAACATTGTT 516
153 IleLeuAspSerLysValAspTyrArgIlyTyrGlnSerMetAspAsp 169
|||||
517 AATTGGACAGCAAGTGATTAAGAGTTATCAATCATCGAGCAACTT 566
169 eIleLysLysAsnThrProGlnGlyPheLysGlySerSerPheLysThr 186
|||||
567 TATTAAAAAACAACACTCCACAAGTTTCAAGCATCAAGTTTAAACTG 616
186 aIleGluValAsnArgLysGluGluValAlaLeuIleMetAsnSerSer 202
|||||
617 TAGAAGTTAAACGCAAGAAAGTTGCTTTATATGAAACCTTCGGGT 666
203 SerThrGlyLeuProLysGlyValGluLeuThrHisGlyAsnIleValTh 219
|||||
667 TCACCGGTTTCCAAAAAGGTGTGCACTTACTCATGAAAAATTTGGTCA 716

```

```

219 rArgPheSerHisAlaArgAspProIleTyrGlyAsnGluValSerPro 236
|||||
717 GGGTTTTCCTCAGCGTAGAGATCCAAATTATGAAACCAAGTTTCACAG 766
236 lYThrAlaIleLeuThrValAlaProPheHisHisGlyPheGlyMetHe 252
|||||
767 GCACGGCTATTTAACTGATGATCAATTCACATGATGTTGGTATGTTT 816
253 ThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLeuThr 269
|||||
817 ACTACTTAGGCTATCTAATCTGTTGTTTTCGATTTGATTCATCTTAACA 866
269 sPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSer 286
|||||
867 ATTTGACGAAGAGACTTTTAAAAACACTGCAAGATTAACAATGTCGAA 916
286 eYValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerLeu 302
|||||
917 GCGTTATCTGTACCGACTTGTGCAATTCCTTAATAGAACTGAATTA 966
303 LeuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyVal 319
|||||
967 CTCGATTAATATGATTTATCAATTTAGTTGAATTCATCTGGCGGAGC 1016
319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeu 336
|||||
1017 ACCTTTATCTAAAGAAATGGTCAACCTGTCTAGACCTTTTAAATTTAC 1066
336 roGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAlaIle 352
|||||
1067 CGGGTGTGTCAGGCTATGTTTAAACAGAAACAACCTCTCAATTAAT 1116
353 IleThrProGluGlyAspAspLysProGlyAlaSerGlyLysValAla 369
|||||
1117 ATCACACCGGAAGCGATGATTAACAGAGTCTTGCGCAAGTTGTGCC 1166
369 oLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGly 386
|||||
1167 ATTATTAAAGCAAGAAATATGATCTTGTATCTAATAAAAACTTGGGCC 1216
386 rAsnArgArgGlyGlyValAlaCysValLysGlyProMetLeuMetLys 402
|||||
1217 CGACACAGCTGCAAGAAATTTGTTAAAGGCTCTTATGTTAAGAAAGT 1266
403 TyrValAspAsnProGluAlaThrArgGluIleIleAspGluGlyTr 419
|||||
1267 TATGTAGATATATCCAGAACCAACAGAAATCATGATGAAGAAAGTTG 1316
419 pLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysPhePhe 436
|||||
1317 GTTGCACACAGAGAAATATTTGGTATTACGATGAAGAAACATTTCTTTA 1366
436 lValAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
|||||
1367 TCGTGGATCTGTTGAAGTCTTTAATCAAAATCAAAAGATTAAGAACCA 1416
452 ProAlaGluLeuGluSerValIleLeuGlnHisProAsnIlePheAsp 469
|||||
1417 CCGCTGAATTAAGATCTGTTCTTTTCAACATCAAAATATTTTATGTC 1466
469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyVal 486
|||||
1467 CGGGCTGTGGCTGCTCCAGATCCATATAGTGTGAGACTTCGGGAGCTG 1516
486 aIleValIleLeuLysGlyLysSerMetThrGluLysGlyValMetAsp 502
|||||
1517 TTGTGTACTTGAAGAAAGAAATCTATGACTGAAAAAGAAATATGAT 1566
503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValAla 519
|||||
1567 TACGTGTAGTCAAGTTTCAAAATGCAAAACGTTGGCTGTGTGTCGC 1616
519 gpHeValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLys 536

```



```

|||||
567 TATTAAAAAACACATCCACAAAGTTTCAAGATCAAGTTTAAACAG 616
186 aLgIuValaInaArgLysGluInValAlaLeuIleMetLnsSerSergly 202
617 TAGAAGTTAAACCCCAAGAACAGAGTGTCTTAAATGAACCTCTTGCGGT 666
203 SerThGlyLeuProLysGlyValGlnLeuThrHisGluasnIleValTh 219
667 TCACACGGTTTGGCAAAAGGTGTGCAACTTACTCATGAAATTTGGTCCAC 716
219 rAngPheSerHisAlaArgAspProIleTyrGlyasnGlnValSerProG 236
717 GCGTTTCTCACGCTAGAGATCCAAATTTATGAAACCAAGTTTCAACAG 766
236 lYThraIleLeuThrValAlaProPheHisHisGlyPheGlyMetPhe 252
767 GCACGGCTATTTTAACTGATAGTACCTTCATCATGCTTTTGGTATGTT 816
253 ThrThrLeuGlyTyrLeuThrCysGlyPheArgIleValMetLeuThrly 269
817 ACTACTTATAGGCTATCTACTGTGTGTTTCCATTTGTCATGTATACGAA 866
269 sPheAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSers 286
867 ATTGACGAGAGACTTTTAAAAACCTGCAGATTCACAAATGTTCAA 916
286 eValIleLeuValProThrLeuPheAlaIleLeuAsnArgSerglyLeu 302
917 GCGTATTTCTGTACCGACTTGTTCGAATTCCTAATAGAAGTCAATTA 966
303 leuAspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyAla 319
967 CTGCAATAATATGATTATCAAAATTTACTGAAATTCATCTGGCGGAGC 1016
319 aProLeuSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeuP 336
1017 ACCTTATCTAAAGAAATTTGGTGAAGCTGTGTGCTAGAGCTTTAATTAC 1066
336 roGlyValArgGlnGlyTyrGlyLeuThrGluThrSerAlaIleIle 352
1067 CCGGCTGTCGCAAGGCTATGTTTAAACGAAACACCTCGCAATTAAT 1116
353 lIeThrProGluGlyAspAspLysProGlyValaSerGlyLysValAlaPr 369
1117 ATCAACCGGAGGCGATGATTAACCAAGAGCTCTCTGCAAGTGTGCC 1166
369 oLeuPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyP 386
1167 ATTATTTAAACAAAGTTATCGATCTTGATCTAATAAAAAAATTGGGCC 1216
386 roAsnArgArgGlyGluValCysValLysGlyProMetLysMetLysGly 402
1217 CGAACACACGTGGAGAGTGTGTGTAAGGGCTCTATGCTTATGAAGGT 1266
403 TyrValAspAsnProGluAlaThrArgGluIleIleAspGluGluGlyTr 419
1267 TATGTAGATATATCCAGAACCAAGAGAAATCATAGTCAAGAGGTTG 1316
419 rLeuHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePheI 436
1317 GTTGCAACAGAGAGATTTGGTATTAAGATGAAGAAAAAATTTCTTTA 1366
436 leValAspArgLysSerLeuIleLysTyrLysGlyTyrGlnValPro 452
1367 TCGTGGATCTTTGAAGCTTTTAACTCAATATCAAGAGATTCAGATCA 1416
453 ProAlaGluLeuGluSerValLeuLeuGlnHisProAsnIlePheAspAl 469
1417 CCGGCTGAATTAAGATCTGTTCTTGGCAACATCCAAATATTTTATGTC 1466
469 aGlyValAlaGlyValProAspProIleAlaGlyGluLeuProGlyAlaV 486
|||||

```

```

1467 CGCGCTCTGGCGTTCCAGATCCTATAGCTGTGAGCTTCGGGAGCTG 1516
486 aLValaLeuLysLysGlyLysSerMetThrGluLysGluValMetAsp 502
1517 TTGTGTACTTGAAGAAAGAAATCTATAGCTGAAAAAGAAATATGGAT 1366
503 TyrValAlaSerGlnValSerAsnAlaLysArgLeuArgGlyValaAr 519
1567 TACGTGTCTAGTCAAGTTTCAAAATGCAAAACGTTTGGCTGGTGTCCG 1616
519 gPheValAspGluValProLysGlyLeuThrGlyLysIleAspGlyLysa 536
1617 TTTGTGGACCAATACCTAAGGTCTCAGCTGGTAAATTAATGAGTAAAG 1666
536 lAlleArgGluIleLeuLysLysProValAlaLysMet 548
1667 CAATTAAGAAATACTAGAGAAACCACTGCTTAAGATG 1704

seq_name: em_pat:E10169
seq_documentation_block:
ID E10169 standard; DNA; UNC; 2019 BP.
XX
AC E10169;
XX
SV E10169.1
XX
DT 08-OCT-1997 (rel. 52, Created)
DT 02-SEP-2000 (rel. 65, last updated, Version 2)
XX
DE DNA encoding streptoavidin-firefly luciferase fusion protein.
XX JP 1995289264 -A/1.
XX
OS unidentified
OC unclassified.
XX
RN [1]
RP 1-2019
RA Tatsumi H., Fukuda M.;
RT "VARIANT TYPE STREPTAVIDIN GENE, FUSION PROTEIN GENE OF VARIANT TYPE
STREPTAVIDIN-FIREFLY LUCIFERASE, NEW RECOMBINANT AND PRODUCTION OF FUSION
PROTEIN OF VARIANT TYPE STREPTAVIDIN-FIREFLY LUCIFERASE",
RL Patent number JP1995289264 -A/1, 07-NOV-1995.
XX KIKKOMAN CORP.
XX
OS None
OC Artificial sequences.
CC PN JP 1995289264 -A/1
CC PD 07-NOV-1995
CC PF 27-APR-1994 JP 1994090275
CC PI TATSUMI HIROKI, FUKUDA MASARU
CC PC C12N15/09, C07K14/36, C07K19/00, C12P21/02, G01N33/53, C12P21/02,
CC PC C12R1:19);
CC CC strandness: Double;
CC CC topology: Linear;
CC CC hypothetical: No;
CC CC anti-sense: No;
CC CC Key location/Qualifiers
CC FH source 1..2019
CC FT CDS 1..2019
CC FT /product="streptoavidin-firefly luciferase
fusion protein"
XX
FH Key location/Qualifiers
FH source 1..2019
FT /db_xref="taxon:32644"
FT /organism="unidentified"
XX
SQ Sequence 2019 BP; 590 A; 412 C; 467 G; 550 T; 0 other:

```

alignment_scores:		546
Quality:	2799.00	
Ratio:	5.126	Gaps: 0
Percent Similarity:	100.000	Percent Identity: 99.267

alignment_block:
nc=00-591-241-6 x F10169

Align seg 1/1 to: E10169 from: 1 to: 2019

3	ASnMeSLuSaNsApGLuAsnLIeValTYrGlyProGluProPheThyP	19
4	
382	TCCTCGGAACGATGAAATAATTGTGTATGTCTCTGAACCACTTTTACC	431
19	oileGluGlySerAlaGlyAlaGlnLeuAtrGlySerMetAspArgT	36
432	TATTGAAGAGGAGTCTGCTGGAGCCAACTTGGCGCAAGTATATGATCGAT	481
36	YrAlaIaIeGluGlyAlaIleAlaPheThrAsnAlaLeuThrGlyAlaAsp	52
482	ATGCAAAACTTGGACCAATTGCTTTTACTAACGCCATTACCGGTGTGCAAT	531
53	TYrThrTYrAlaGluTYrIeGluGlyIleuSerCysCysLeuGlyGlnAlaLe	69
532	TATAGCTACGGCGAATACTTACAAAAATCATCTGCTGCTGAGAGAGCTTT	581
69	uIyAsnIlyGlyLeuValAlaAspGlyArgIleAlaIleuCysSerGlnA	86
582	AAAGAAATTAAGTTTGTTGGTGTGATGAGAAAGAAATTCGTTATCATCAGTAA	631
86	sNCysGIuGluPhePheIleProValIleuAlaGlyLeuPheIleGlyAla	102
632	ACTGTGAAGAGTTCCTTATTCCTGTATATTAGCCGGTTATTATAGTGATC	681
103	GlyValAlaIleProThrAsnGluIleTYrThrLeuAtrGluLeuValHisSe	119
682	GGTGGGGCTCCCAACATGATGATTTACACTACCTACGGAATGGTTACAG	731
119	IleuGlyIleSerIysProThrIleValPheSerSerIysGlyLeuA	136
732	TTTAGGCATCTCTAAGCCCAACAATTGATTAGTTCTAAAAAGAAATTAG	781
136	sPlyValIleThrValGlnIysThrValThAlaIleIysThrIleVal	152
782	ATTAAGATTATTACTGTACAAAAACGGTAACGCTATTTAAACCACTGGTT	831
153	IleIeuAspSerIysValAspIlyrArgIlyTYrGlnSerMetAspAsnPh	169
832	AAATTGGACAGCAAACTGGATTATAGAGGTTATCAATCCATGCACAACCT	881
169	eIleIysIysAsnThrProGlnIyPheIysGlySerSerPheIysThyV	186
882	TYTTAAAAAAACACTCCCAAGAGTTTCAAGAGATCAAGTTTAAAAACTG	931
186	AlGluValAsnArgIysGluGlnValAlaIleuIleMetAsnSerSerGly	202
932	TAGAACTTAACCGCAAGAACCAAGTTGCTCTTATTAATGAACCTCTTGCGGT	981
203	SeTrThGIyLeuProIyGlyAlaGlnLeuThrHisGluAsnIleValTh	219
982	TCAACCGGTTTGGCAAAAGGTGTGCAACTTACATGAAAAATTGGTGCAC	1031
219	rArgPheSerHisAlaArgAspProIleTYrGlyAsnGlnValSerProG	236
1032	GGCTTTTTCACAGTCAGAGATCCAAATTATGGAACCAAGTTTCACACAG	1081
236	IyThrAlaIleLeuThrValValProPheHisIshGlyPheGlyMetPhe	252
1082	GCACGGCTATTTTAACTGATGATACCATTCATCATGTTTGGATGCTTT	1131
253	ThrThleuGlyTYrIeuthrCysGlyPheArgIleValMetIeuthrIly	265

1132	ACTACTTAGAGCTATCTA	CTCACTGGTGGTTTGGTATGTCATGTTAACGA	1181
269	sPhesApGIuGIuThr	PheLeuLysThrLeuGlnAspTyrLysCysSerS	286
1182	ATTTCGCGAAGAGACTTTT	AAAAACCTCAAGATTACAAATGTTCAA	1231
286	erValIleLeuVal	ProThrLeuPheAlaIleLeuAsnArgSerGIuLeu	302
1232	GGCTATTCTTTACCAGACTT	GTTTGCATTTCTTAATAGAAAGCAATTA	1281
303	LeuAspLysTyrAspLeuSer	AsnLeuValGIuIleAlaSerGIuGIuAl	319
1282	CTTCGATAAATATGATTAT	TACAAATTACTTAATTTGCATCTGGCGGAGC	1331
319	apProLeuSerLysGIuIle	GILGILAlaValAlaArgArgPheAsnLeuP	336
1332	ACCTTTATCTTAAGAATAT	GGTGGAGCGCTTGCTAGACGTTTATTTATTAC	1381
336	roGIuValArgGlnGIuTyr	GIuLeuThrGIuThrSerAlaIleIle	352
1382	CGGCTTGCTGTAAGCTAT	TGGTTTACAGAAACCACTTCGCAATTATT	1431
353	IleThrProGIuGIuAspAsp	LysProGIuAlaSerGIuLysValAlaPr	369
1432	ATACACACCGAAGCGCAT	ATATACACAGAGTCTTGTGCAAAAGTTGTGCC	1481
369	oleuPheLysAlaLysValI	leAspLeuAspThrLysLysThrLeuGIuP	386
1482	ATTATTAAAGCAAAAGTAT	GCATCTTGATCTATAAAAACTTGGGCGC	1531
386	roAsnArgArgGIuGIuVal	LysValLysGIuProMetLeuMetLysGIu	402
1532	CGAACACAGACTGGAGAA	GTGTTGGTAAAGGGCTCATGCTTATGAAGGT	1581
403	TyrValAspAsnProGIuAl	aThrArgIuIleIleAspGIuGIuGIuTyr	419
1582	TATGTGATATATCCAGAG	CAACAAAGAAATCATATGTAAGAAAGGTGTG	1631
419	pleuHisThrGIuAspIle	GILGIuTyrThrAspGIuGIuLysHisPhePheI	436
1632	GTTCACACAGGAGATATT	GGGATTATGCATAGAGAAACCACTTCTTTA	1681
436	IleValAspArgLeuLysSer	LeuIleLysTyrLysGIuTyrGIuValPro	452
1682	TCGTGGATTCGTTTGAAG	TCTTATCAATATACAAAGATATCAATGATCA	1731
453	ProAlaGIuLeuGIuSerVal	LeuLeuGlnHisProAsnIlePheAspAl	469
1732	CTGCCTCAATTAAATCTG	TTCTTTGGCAACATCCAAATATTTTGGATGC	1781
469	agIValAlaGIuValProAsp	ProIleAlaGIuLeuProGIuAlaVal	486
1782	CGGGCTGCTGGCGCTTCA	GATCTTACTGCTGTAGCTTCGGGAGCGTG	1831
486	alValValLeuLysLysGIu	LysSerMetThrGIuLysGIuValMetAsp	502
1832	TTGTGTTGACTTGAAGAAA	AGGAAATCTATACATCGAAGAAACAACTAATG	519
503	TyrValAlaSerGIuValSer	AsnAlaLysArgLeuArgGIuValAlaArg	536
1882	TACCTGTCTATCAAGTTA	ATAATGCAAAACGTTTCCGTGGTGGTGTCCG	1931
519	qPheValAspGIuValPro	LysGIuLeuThrGIuLysIleAspGIuLysA	536
1932	TTTTGTGAGACGAAGTACT	TAAAGTCTCATCTGTATAAATTTGACAGGTAAAG	1981
536	IaIleArgGIuIleLeuLys	LysProValAlaLysMet	548
1982	CAATTAGAGAAATCTCA	AAACCAAGTGTCTAAGATG	2019

OM of: US-09-581-241-4 to: EST: * out_format : pfs
Date: Nov 29, 2001 2:32 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODE=frame-pzn.model -DEV=x1h
-Q/cgn2.1/USPTO-SPool/US09581241/runat-28112001.173311.14853/app-query.fasta.1.1226
-DB=EST -QFMT=fastap -SUFFIX=1st -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -GAPOP=6.000
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DEL0P=6.000
-DELEXT=7.000 -START=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=200000000 -USBR=US09581241 -CGN1.1.5643 -NCPU=6
-ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPHY
-WAIT -THREAS=1

Search information block:

Query: US-09-581-241-4
Query length: 548
Database: EST: *
Database sequences: 11351937
Database length: 1077921985
Search time (sec): 1937.630000

Score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Document
gb_gss:CNS01FCA	+	623.00	1135.63	2.0e-57	975	AL141659 Anopheles gambiae GSS
gb_est1:BG445463	+	552.00	1057.72	9.4e-50	882	BG445463 GA_Ea0028K14f Gossyp
gb_est1:BE036528	-	541.00	1034.21	1.9e-48	1066	BE036528 MP01A02 MP Mesembryan
gb_est1:AU214129	-	513.00	983.22	1.3e-45	741	AU214129 AU214129 unpublished c
gb_est2:CG83857	-	512.00	982.41	1.5e-45	664	C83857 C83857 Dictyostelium dis
gb_est2:BF92667	-	507.50	973.32	4.7e-45	683	BF92667 AT01136.3prime AT Dros
gb_est1:AI186799	+	505.50	969.65	7.6e-45	667	AI186799 EST245121 tomato ovary
gb_est2:BG14540	+	504.50	967.86	9.1e-45	876	BG14540 HVSMMK002N03f Hordeum
gb_est1:AM691106	+	504.50	967.86	9.1e-45	876	AM691106 NF041C08ST1F1000 Davel
gb_est1:AI485586	+	482.50	923.32	2.2e-42	625	AI485586 EST243907 tomato ovary
gb_est1:BE454671	+	482.00	923.36	2.9e-42	686	BE454671 HVSMMK0094P16f Hordeum
gb_est2:BF623201	+	481.50	920.55	4.1e-42	817	BF623201 HVSMMK0015K22f Hordeum
gb_est2:BF064865	-	479.00	915.80	7.6e-42	806	BF064865 HV_CBP002D1D2f Hordeum
gb_est2:BE820931	-	477.50	913.32	1.0e-41	772	BE820931 GM700013A20B2 GM-r107C
gb_est1:AV558307	+	474.50	910.02	1.6e-41	604	AV558307 Arabidopsis t
gb_est1:AM329421	+	472.50	906.35	2.5e-41	590	AM329421 N200656e rootphos(-) M
gb_est1:AU214473	-	472.50	905.53	2.8e-41	638	AU214473 AU214473 unpublished c
gb_est1:BE555239	-	469.50	901.35	4.8e-41	543	BE555239 SP87410.Y1 Gm-cl045 G1
gb_gss:CNS01MDU	-	468.00	897.25	1.6e-40	977	AL149424 Anopheles gambiae GSS
gb_gss:CNS01LBB	-	465.50	890.85	1.9e-40	769	BI1311228 EST312979 GESD Medica
gb_est2:BI311229	+	465.50	890.83	1.9e-40	693	BI311228 EST312979 GESD Medica
gb_est1:AM621420	+	464.00	889.00	3.0e-40	582	AM621420 EST312218 tomato root
gb_est1:AA979869	+	464.00	887.00	3.0e-40	764	AA979869 MEST2-F3.TW1412.Seq 19
gb_est1:AV680327	-	463.50	889.32	2.3e-40	558	AV680327 AV680327 Nori Satoh un
gb_est1:AU219684	-	458.50	878.70	8.8e-40	604	AU219684 AU219684 unpublished c
gb_est2:BG50544	+	456.00	874.47	1.5e-39	567	BG50544 EST482271 CYN Medicag
gb_est1:AM218471	+	454.50	871.50	2.2e-39	569	AM218471 EST303654 tomato radic
gb_est1:AM560940	+	453.50	866.07	3.5e-39	655	AM560940 EST315988 D5IR Medica
gb_est2:BF200141	+	451.50	863.99	7.3e-39	595	BF200141 WHE2252.E02.104ZE Trif
gb_est1:AM329094	+	450.00	862.22	7.3e-39	623	AM329094 NF052809PT1F1076 Droug
gb_est2:BF634495	+	449.00	861.74	7.8e-39	644	BF634495 EST50823 GPOD Medica
gb_est2:BI309413	+	447.50	856.50	1.5e-38	644	BI309413 EST50823 GPOD Medica
gb_est2:BE597274	+	445.50	852.59	2.5e-38	644	BE597274 P11_70.C05.g1.A002 Pat
gb_est1:BE597274	+	445.00	853.37	2.0e-38	495	AL284477 4A3B-AAX-B-04 F Anophe
gb_est1:AL3398862	-	443.00	849.67	3.7e-38	535	AL3398862 MCB57H09F1 MCB Medica
gb_est1:AM776080	+	443.00	849.01	4.5e-38	625	AM776080 EST335145 D5IL Medica
gb_est1:AI488821	+	438.00	840.11	1.2e-37	522	AI488821 EST247160 tomato ovary
gb_est1:BE330067	+	437.50	838.30	1.6e-37	565	BE330067 s073505.Y1 Gm-cl040 G1
gb_est2:BE803370	+	435.50	833.10	3.1e-37	639	BE803370 s755603.Y1 Gm-cl051 G1
gb_est2:BI425803	+	434.50	833.75	2.8e-37	498	BI425803 sah72103.Y1 Gm-cl049 G

seq_name: gb_gss:CNS01FCA
seq_documentation block:
LOCUS CNS01FCA 975 bp DNA GSS 01-JUN-2001
DEFINITION Anopheles gambiae GSS SP6 end of clone 04D06 of NotreDame1 library
from strain PEST of Anopheles gambiae (African malaria mosquito),
genomic survey sequence.
ACCESSION AL141659
VERSION AL141659.1 GI:6999777
KEYWORDS GSS.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
Culicoidae; Anopheles.
1 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 975)
REFERENCE
1 (bases 1 to 975)
GENOSCOPE.
Direct Submission
Submitted

```

223 sAlaArgSPProLeuThrLeuGlnValSerProGlyThrAlaLeu 240
284 CTGGAAGAGTCGGCTAAGCTGCTGGAGCTACCGGATCAGCTGGTACAC 333
240 eutThrValProPheHisHis.....GlyPheGlyMetPheThr 253
334 TACAGAGTACCTCTGCTGCTCCAGCTGCGCGGTGCGGCTG..... 377
254 ThrLeuGlyThrLeuPheGlyPheArgGlyLeuMetLeuThrLysPhe 270
378 ...CTGAATATGGTTACCAACAATGCCGCTGTGTGTGATGCCACGATT 424
270 eAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSerSer 287
425 TGATTCACACCTGTTCTCTAGCTGTATGAAAGTACGGGTAAATCTAA 474
287 aAlaLeuValProThrLeuPheAlaLeuLeuAsnArgSerGluLeuLeu 303
475 TGACACTGCTACCTCCGCTGATGCTCTTCTGTGCAAAACATCCATGCTG 524
304 AspLysTyrAspLeuSerAsnLeuValGluLeuAlaSerGlyGlyAlaPro 320
525 GAAACTACATCTCTCTCTGCTGATGACCTGTTCTGTGCGCGGCTCC 574
320 oLeuSerLysGluLeuGlyGluAlaValAlaArgArgPheAsnLeuProG 337
575 GTTGACCAAGAGATCGAAGATCAGTGGCGGAGCGACTGGTGTAGCTT 624
337 LValArgGlnGlyTyrGlyLeuThrGluThrPheSerAlaLeuLeu 353
625 TCATTGGCAAGGATGAGCATGATGAAACCAAGCTCGGTGTGTGATG 674
354 ThrProGluGlyAspAspLysProGlyAlaSerGlyLysValAlaProLe 370
675 CAGAGCGAGTTCGAAACAAGCGGCGAGCTGGGAAGGTGGCGCATGGC 724
370 uPheLysAlaLysValLLeuAspLeuAspThrLysLysThrLeuGlyPro 387
725 CCAGTGGGGAAGGTGATGACCCGCGAGCGGCGCTGCGCTGCGACCGA 774
387 sNArgArgGlyGluValLysValLysGlyProMetLeuMetLysGlyTyr 403
775 ACCAGCGGGGTGAGCTGCTTTAAGGCTCATGTGATGAAAGGATAT 824
404 ValAspAsnProGluAlaThrArgGluLeuLeuAspGluGlyTyrPhe 420
825 GTGGCGCAAGACGACGCC.....ATCGATGCTGACGGTTGGCT 862
420 uHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePheIle 437
863 GCACAGGGGCGACATTGCTGCTACTAGCATGATGATCAGAGTTTTCATG 912
437 aAlaAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValProPro 453
913 TGGACCGGATCAAGAGCTGATCAAGTACAGGCGCTTCARGTACCGCG 962
454 AlaGluLeuGlu 457
963 GCCGAGCTGGAA 974

seq_name: gb_est2.BG445463

seq_documentation_block:
LOCUS BG445463 882 bp mRNA EST 15-MAR-2001
DEFINITION GA_Ea0028K14f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboresum cDNA clone GA_Ea0028K14f, mRNA sequence.
ACCESSION BG445463
VERSION BG445463.1 GI:13355211
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 882)
Wing, R.A., Fritsch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry
D., Wood, T.C., Leslie, A. and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
CONTACT: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAACTACCTCATTATAGG
High quality sequence stop: 733.
Location/Qualifiers
1..882
/organism="Gossypium arboreum"
/strain="AKA"
/db_xref="taxon:29729"
/clone="GA_Ea0028K14f"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 250 a 173 c 226 t 1 others
ORIGIN

alignment_scores:
Quality: 552.00 Length: 277
Ratio: 2.816 Gaps: 4
Percent Similarity: 70.758 Percent Identity: 41.155

alignment_block:
US-09-581-241-4 x BG445463 ..

Align seg 1/1 to: BG445463 from: 1 to: 882
263 ArgIleValMetLeuThrLysPheAspGluGluThrPheLeuLysThrLe 279
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 CGAGTTTATCATCGAAGAGTTCGACATCGGCTCTGTTGAGCTAAT 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
279 uGlnAspTyrLysCysSerSerValIleLeuValProThrLeuPheAla 296
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
56 ACAGAAATACAAAGTAAACATTGCTCCGATGTCACCGCATGATTGCG 105
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
296 LeuLeuAsnArgSerGluLeuLeuAspLysTyrAspLeuSerAsnLeuVal 312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 CCATCCCTAAGTATCGAAGAACTGAAATTCAGATTGTCATCGGAGG 155
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
313 GlnIleAlaSerGlyGlyAlaProLeuSerLysGluIleGlyGluAlaVal 329
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
156 ATGTGAAGTCCGTCGCCCTCGTTGGGTCAAGAGCTTGAGAGATGCTGT 205
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
329 LAlaArgArgPheAsnLeuProGlyValArg.....GlnGlyTyrGlyL 344
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 AAAAGTCAAGTTT.....CTGTGTGCCAAATTTGGACAGGGTTATGAA 249
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
344 eutThrGluThrSerAlaIleIleIle.....ThrPro 355
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
250 TCACAGAAAGCTGACACAGTTCAGCAATGTGTTGGATTGCCAAGGAA 299
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
356 GluGlyAspAspLysProGlyAlaSerGlyLysValAlaProLeuPheLys 372
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
300 CCTTTGAATGAAATCGGAGACTGTGGAGCTGTTGTAAGAACGCGGA 349
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
372 sAlaLysValIleAspLeuAspThrLysLysThrLeuGlyProAsnArg 389
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

350 GATGAAATGCTGACCCCGAAACCGGTTTCTTCCGCCAGAACGAG 399
389 rgglgluValuAluValysglPrometleuMetlYsglYrValasp 405
400 CCGGAGGATTTGCATTCGAGAGATCATGATCAAGAGATACCTTAAT 449
406 AsnProgluAlaThrArgluIlelleAspGluIuLyrThreHsth 422
450 GACCTTGAGGCTGCTGAGCACCATTGACAAAGATGCTGTTCATAC 499
422 rglYAspIleGlyTrpYrAspGluIuLysHisPhePheIleValasp 439
500 CGGAGACATTTGTTACATTCAGACGATGATGAACTCTTCATCTGATC 549
439 rglYLeuYsSerleuIleLysTrpYsglYrGluValProProAlaGlu 455
550 GATTCAGAAATGATTAATAAATACAAAGGTTTCAAGTTCCTCCGCGAG 599
456 leuGluSerValleuLeuGlnHisProAsnIlePheAspAlaGluVal 472
600 CTCGANNACGATCTCATTTCTACCTGAAATCATCGATGCTGCTGTGT 649
472 aglYValProAspProIleAlaGluLeuProglYAlaValVal 489
650 CCGCATGAGATGAGACAGCTGGGAGAGTACCTGTTGATTTGGTGA 699
489 euLysLysGlyLysSerMetThrGluLysGluValMetAspTrpVala 505
700 GATCAGATAAATCTCAGATCAATGAAATGAAATCAACCAATATATTCG 749
506 SerGluValSerAsnAlaLysArgleuArgGlyValArgPheValas 522
750 AAACGGGGTGGTCTTATAGGAGATAGCCGGGGGTTTC...TTCATTGG 796
522 polYValProLysGlyLeuThrGlyLysIle 532
797 AGCCCTTCCAAAGGCCCATCAGGGAAGATC 827

seq_name: gb_est1:BE036528

seq_documentation_block:
LOCUS BE036528 1066 bp mRNA EST 07-JUN-2000
DEFINITION MP01A02 MP Mesembryanthemum crystallinum cDNA 5' similar to
commenrate:coenzyme a ligase, mRNA sequence.
ACCESSION BE036528
VERSION BE036528.1 GI:8331537
KEYWORDS EST.
SOURCE common ice plant.
ORGANISM Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllidae; Caryophyllales; Alzooceae; Mesembryanthemum.
1 (bases 1 to 1066)
Bohnerl,H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M., Ferrea
,H., Kawasaki,S., McColough,A., Michalowski,C.B., Palacio,C.,
Scara,G., Wheeler,M. and Zepeda,G.R.
Functional Genomics of Plant Stress Tolerance
Unpublished (2000)
Contact: Michalowski,C.B.
University of Arizona
Bio Sciences West room 513, Tucson, AZ 85721, USA
Tel: 520-621-7982
Fax: 520-621-1697
Email: chm@u.arizona.edu
An open reading frame exists.
Insert Length: 1 Std Error: 0.00.

FEATURES
Source
1..1066
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone_lib="MP"
/tissue_type="apical meristem and leaf primordia"
/dev_stage="6 weeks"

```

```

BASE COUNT 265 a 220 c 279 g 294 t 8 others
ORIGIN

alignment_scores:
Quality: 541.00 Length: 295
Ratio: 2.678 Gaps: 5
Percent Similarity: 68.475 Percent Identity: 40.339

alignment_block:
US-09-581-241-4 x BE036528 ..
Align seg 1/1 to: BE036528 from: 1 to: 1066

218 ValThrArgPhe.....SerHisAlaArgAspProIleTrpGlyAsnGlu 232
4 ATTAACGGCTATAGTGATGCATATATAAATCTCCG..... 39
232 nvalSerProGlyThrAlaIleuThrValValProPheHisGlyP 249
40 .....CCAAATGGCGCTCGTTCTTCTCCGCTGTTGATGCTCT 76
249 heGlyMetPheThrThrLeuGlyTrpLeuThrCysGlyPheArgIleVal 265
77 TTGGTTCTTCATGCTGCTGAGAGCGGTGCTGGGCCGAATCATGATGTT 126
266 MetLeuThrLysPheAspGluIuThrPheLeuLysThrLeuGlnAsp 282
127 GTATATGCAAGGTTGATTCGAGAACATGTGTGAGCGCTGGAGAGAT 176
282 rlyCysSerSerValIleLeuValProThrLeuPheAlaIleLeuAsn 299
177 TCGATTACATACATTCGCGTTTCGCCCATTTGGTGTGCGCATGACAA 226
299 rglYSerGluLeuLysPlyLysTrpYrAspLeuSerAsnLeuValGluIleAla 315
227 AATCGAGACTTACAGATAGATAGATCTCAGACCGCTGAAGTCTCGCT 276
316 SerGlyGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArg 332
277 TGTGGTGGCGCTCTCTTGGGAAGAGGTGCAACGGGTTTGAGACTCG 326
332 gPheAsnLeuProGlyValArgGluGlyTrpGlyLeuThrGluThrLys 349
327 CTTCGCCAGCTGAGATAGATAGATAGATAGATAGATAGATAGATAG 376
349 eraAlaIleIleIleThr.....ProGluLysAspAspLysProGlyAla 363
377 GCGCGGCTGCAAGAGACGTAGGCGCTGAGGAATCTGAAACATATGTTCT 426
364 SerGlyLysValValProLeuPheLysAlaLysValIleAspLeuAsp 380
427 GTTGGCGGCTTACAGAGCTTGTGGAAGCAAGATATGATTCGCTAAC 476
380 rlyLysThrLeuGlyProAsnArgArgGlyGluValCysValLysGlyP 397
477 CGGAGAGGCTTACACCTGCTGTAAGAGGGGAGCTGCTGCTGGTGGC 526
397 rometleuMetLysGlyTrpValAspAsnProgluAlaThrArgGluIle 413
527 CAACCATATGATAAAGCATATGCTGGAGATAGATAGATAGATAGATAG 576
414 IleAspGluGluGlyTrpLeuHisThrGlyAspIleGlyTrpYrAsp 430
577 NTCCATCCANATGATGCTTGAAGACTGAGATGCTTTGTTATTTTGCATC 626
430 uGluLysHisPhePheIleValAspArgLeuLysSerleuIleLysTrp 447
627 TGATGGCTTCTTCATATGTTGATTCGATTAAGGAATNGATTAAGATAC 676
447 ysglyTrpGluValProProAlaGluLeuGluSerValleuGlnHis 463

```

```

677 AGGATTCACNGTCCCTCCAGCTGAGTTGGAGCGGTTCATCACTATCCCAT 726
464 ProasnlIerheapsAlagIyValAlaGIyValProAspProllealagl 480
|||||:||||| ||||| :||| |||||
727 CCAGATAATGCTGTAATGCTGCCTGTGTCCTGATCCAGATCAGAACGCG 776
480 yGlulnPrGclYAlaValAlaVallleuLyslySGLy.LysSerMetThr 496
|||||:||||| ||| ||||| ||||| :|||
777 ACAAAATACCTATGCCATTGTGTGTC..AGGAATNCAGAAAGCACCATTTTC 823
497 GlulysGlulValMetAspyrYValAlaISergln 507
:::|||||:||||| ||||| ||||| :|||
824 TGGCTCCACCTCATGATATGATGCTAACAAGC 856

seq_name: gb_est1:AU214129

seq_documentation_block:
LOCUS AU214129 741 bp mRNA EST 17-JUL-2001
DEFINITION AU214129 unpublished oligo-capped cDNA library, stage L2
Caenorhabditis elegans cDNA clone YK809f05 3', mRNA sequence.
ACCESSION AU214129
VERSION AU214129.1 GI:14852286
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea
1 (bases 1 to 741)
Koehara,Y., Shih-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
A complementary view of the C.elegans genome
Unpublished (2001)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
source
1..741
Location/Qualifiers
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk809f05"
/cGene_lib="unpublished oligo-capped cDNA library, stage
L2"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L2"
BASE COUNT 216 a 158 c 125 g 241 t 1 others
ORIGIN

alignment_scores:
Quality: 513.00 Length: 223
Ratio: 3.090 Gaps: 5
Percent Similarity: 74.439 Percent Identity: 47.982

alignment_block:
US-09-581-241-4 x AU214129/rev ..

Align seg 1/1 to reverse of: AU214129 from: 1 to: 741

320 ProteuserylGluInleGlyGluAlaValAlaArgaYPHe..AsnLe 335
|||||:|||||:|||||:|||||:|||||:|||||
739 CCAGCTGGAAAGCAATTCATGAGAATTCGAAGAGAGACTATACAAATTT 690
335 uProGlyValARGgInGLyTYrGYlVeUthrGrlgnThrsEraIaleI 352
|-::::::::::|||:|||||:|||||:|||||:|||||
689 GAATATACATCCAACAAGATATGATGATGCGAATCTTCATGCTAGTC 640
352 lelleThProGluclYaspAsplyPro...glYAlaSerelyLysVal 367
```

seq_name: gb_est2:C83857

seq_documentation_block:

LOCUS C83857 664 bp mRNA EST 28-APR-1999

DEFINITION C83857 Dictyostelium discoideum SS (H. Urushihara) Dictyostelium

ACCESSION C83857

VERSION C83857

KEYWORDS

SOURCE EST.

ORGANISM Dictyostelium discoideum.

ORGANISM Dictyostelium discoideum

REFERENCE 1 (bases 1 to 664)

AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M., Yoshino,R., Mitera,B.N., Pl.M., Sato,T., Takemoto,K., Yasukawa,H., Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.

TITLE The Dictyostelium developmental cDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development

JOURNAL DNA Res. 5 (6), 335-340 (1998)

MEDLINE 99156227

COMMENT Contact: Hideko Urushihara

Institute of Biological Sciences

University of Tsukuba

3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan

Email: d402hu@sakura.cc.tsukuba.ac.jp

PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
POLYA-No.

FEATURES
Source Location/Qualifiers
1.664
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone_lib="SSA391"
/dev_stage="slug"
/clone_stage="slug"

BASE COUNT 234 a 105 c 124 g 201 t
ORIGIN

alignment_scores:
Quality: 512.00 Length: 200
Ratio: 3.261 Gaps: 0
Percent Similarity: 78.500 Percent Identity: 47.500

alignment_block:
US-09-581-241-4 x C83857 ..

Align seg 1/1 to: C83857 from: 1 to: 664

```

340 GlnGlyTyrGlyLeuThrGluThrSerAlaIleIleIleThrProG1 356
|||||
2 CAGGTTATGGTGTCCGACCACTAGTCCAGCTTTTGTATTCAG 51
|||||
356 uGlyAspAspLysProGlyAlaSerGlyLysValAlaProLeuPheLysA 373
|||||
52 TGGATTAGTAAATCTGGTCCGCTGATCCCTTACCACATCAATGG 101
|||||
373 lAlaValIleLeuPheLysPheLysThrLeuGlyProAsnArgArg 389
|||||
102 CTAATATCATCTCTCCAGAGACTGTGGAATTTAGTATGGGCAAAAG 151
|||||
390 GlyGluValCysValLysGlyProMetLeuMetLysGlyTyrValAspAs 406
|||||
152 GGTGAATCTGTATTAAGGTCCAATGTATGTGGTTATTAACA 201
|||||
406 nProGluAlaThrArgGluIleIleAspGluGluGlyTrpLeuHisThrG 423
|||||
202 TCAAAAGGCAACCAATGAAGTCATGAATGAATGATTTAAAAACTG 251
|||||
423 lAspIleGlyTyrTyrAspGluGluLysHisPhePheIleValAspArg 439
|||||
252 GTCATATTGGTACGTCGATGAAGATGGTATTACTTATCGTCATAGA 301
|||||
440 LeuLysSerLeuIleLysTyrLysGlyTyrGluValProProAlaGluLe 456
|||||
302 TCAAAAGGCAATGATCAATGAAGTTCCAAAGTTCCTCCGCGCAAT 351
|||||
456 uGlySerValIleLeuGluHisProAsnIlePheAspAlaGlyValAlaG 473
|||||
352 GGAGGCAATTCATTAATCATCAAAAGTTCAGATGCTGTGTAGTAG 401
|||||
473 lValProAspProIleAlaGlyGluLeuProGlyAlaValAlaValLeu 489
|||||
402 GTCCTTCAAAAGCTATAGGTGCAAGTCAAGACAGAGCTTCTCTGTATT 451
|||||
490 LysLysGlyLysSerMetThrGluLysGluValMetAspTyrValAlaSe 506
|||||
452 AAAAATAATGATCTCTAACGAAAAAGAACTTGAATGGGCCATCC 501
|||||
506 rGluValSerAsnAlaLysArgLeuArgGlyLysValAlaArgPheValAsp 523
|||||
502 AAAAATTTGCAATTTAATCAATTCAGAGGTGTATCTTTTCATACCTG 551
|||||
523 lValAlaProLysGlyLeuThrGlyLysIleAspGlyLysAlaIleArgGlu 539
|||||
552 CATTTCTTAATCTGCACACCGGTAACCTATTACGTAATAAATCTTAAGAT 601
|||||

```

seq_name: gb_est2:BF492687

seq_documentation_block:

LOCUS BF492687 683 bp mRNA EST 19-Apr-2001
DEFINITION AT01136.3prime AT Drosophila melanogaster adult testes potB7.

Drosophila melanogaster cDNA clone AT01136.3 similar to CG6178:
Pban0006178 enzyme located on: 3R 95E1-95E1.1: 04/07/2001, mRNA
sequence.

ACCESSION BF492687
VERSION BF492687
KEYWORDS EST

SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 683)
Stapleton, M., Broksstein, P., Hong, L., Aghayani, A., Baxter, E., Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C.J., Nuno, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celnik, S. and Rubin, G.M.

TITLE BDGP/HM1 AT Drosophila EST Project
JOURNAL Unpublished (2000)
COMMENT Other ESTs: AT01136.5prime
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd
Pax: 510 486 6798
Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
hit genomic AE003746: arm:3R [19764572..19982668]
estimated-cyto:95D5-95F1: 04/07/2001
Plate: AT.11 row: C column: 12
High quality sequence strop: 631.

FEATURES
source Location/Qualifiers
1..683
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="AT01136"
/clone_lib="AT Drosophila melanogaster adult testes potB7"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DHS-alpha. Plates AT.121-AT.319: DHS-alpha Tona"
/note="Organ: ADULT testes; Vector: potB7; Site:1: EcoRI; Site:2: XhoI; The mRNA for the testis library was made from testes and seminal vesicles hand dissected from 0-3 day old Ore-R males. RNA kindly provided by the lab of Margaret Fuller. Sized fractionated cDNAs were directly ligated into potB7. Plasmid cDNA library."

BASE COUNT 169 a 188 c 149 g 177 t
ORIGIN

alignment_scores:
Quality: 507.50 Length: 176
Ratio: 3.574 Gaps: 1
Percent Similarity: 80.682 Percent Identity: 57.955

alignment_block:
US-09-581-241-4 x BF492687/rev ..

Align seg 1/1 to reverse of: BF492687 from: 1 to: 683

```

373 AlAlaValIleAspLeuAspThrLysLysThrLeuGlyProAsnArg 389
|||||
637 GCACAGGTGATGATCCCGACACCGGCAAGCTATTGGGGCCCAACGAGCG 588
|||||
389 gGlyGluValCysValLysGlyProMetLeuMetLysGlyTyrValAspA 406
|||||
587 CGGCGAGCTTTGTTTAAAGCGACGACATCATGAAGGCGCTACATCGAG 538
|||||

```

```

406 snpGluAlaThrArgGluIleLeuAspGluGluGlyThrLeuHisThr 422
      ::::::::::::::|
537 ATAGAGAGTCCACGACGACCGCCATC...AAGAGCGGTGGTGCATCAT 491
      |||||
423 GUAAspIleGlyTyrAspGluGluHisIshpheIleValAspAsp 439
      |||||
490 GGGCATTTGTGGTACTATGATGATTTGTGATTTCTTCATCGTGCACG 441
      |||
439 GLeuYSerLeuIleLysTyrGlyTyrGlnValProProIleGlu 456
      ::::::::::::::|
440 CATCAAGAGACTGATCAATACAAAGGATACAGAGTCCCGCGCACAGA 391
      |||||
456 euGluSerValLeuLeuGlnHisProAsnIlePheAspAlaGlyValAla 472
      ::::::::::::::|
390 TTGAGGCTCTGCTGCTCACCAAGATTAAGATGCGCGCGCATTT 341
      |||
473 G1ValProAspProIleAlaGlyGluLeuProGlyAlaValValLe 489
      |||
340 GGAAGGCCACACGAGGAGGCTGGCGAGCTGCCGTCGATTTGTCTAA 291
      ::::::::::::::|
489 uLysLysGlyLysSerMetThrGluGluValMetAspTyrValAla 506
      ::::::::::::::|
290 ACAGGCTAATGTCATCACTGACCGAGAACGAAGTCATTCAGTTGTCTA 241
      |||||
506 erGlnValSerAsnAlaLysArgLeuArgGlyValAlaArgPheValAsp 522
      ::::::::::::::|
240 ACACGCGCTCGCCGCGCACGCTTAAGGGGCGCGATCTTGTGTGAC 191
      |||||
523 GUAValProLysGlyLeuThrGlyLysIleAspGlyLysAlaIleArg 539
      |||||
190 GAATTCGCAAGAACCCGACGTCGACGATCTGCGTCGCGGGA 141
      |||||
539 uIleLeuLysLysProValAlaLysMet 548
      ::::::::::::::|
140 AATGCTTAAGAACCAAAATCCAAAGTTG 113

seq_name: gb_est1:A1486799

seq_documentation_block:
LOCUS      A1486799      667 bp      mRNA      EST      18-MAY-2001
DEFINITION EST245121 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
            CLED11D3, mRNA sequence.
ACCESSION  A1486799
VERSION    A1486799.1 GI:4382170
KEYWORDS   EST.
SOURCE     tomato.
           Lycopersicon esculentum
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           Asteridae; eusterids I; Solanales; Solanaceae; Solanum;
           Lycopersicon.
REFERENCE  1 (bases 1 to 667)
AUTHORS   Alcalá, J., Vrebalov, J., White, R., Matern, A. L., Vision, T., Holt, I. E.,
           Liang, F., Upton, J., Romling, C. M., Craven, M. B., Fujii, C. Y., Bowman,
           C. L., Mierman, W., Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley,
           S. D. and Giovannoni, J.
           Generation of ESTs from tomato carpel tissue
           Unpublished (1999)
COMMENT    CUGI
           Clemson University Genomics Institute
           100 Jordan Hall, Clemson, SC 29634, USA
           Email: http://www.genome.clemson.edu/orders/index.html.

FEATURES
     source
         1..667
            /organism="Lycopersicon esculentum"
            /cultivar="TA496"
            /db_xref="taxon:4081"
            /clone="CLED11D3"
            /clone_lib="tomato ovary, TAMU"
            /tissue_type="carpel"
            /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
            /lab_host="X11-Blue MRP"

```

```

/Note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; CLED - Tomato Carpel EST library. OligodT-primed and
directionally cloned cDNA in vector lambda Zap II with 5'
and 3' ends located at the EcoRI and XhoI sites,
respectively."
BASE COUNT      201 a      120 c      155 g      191 t
ORIGIN

alignment_scores:
    quality: 505.50      length: 222
    ratio: 3.159
    percent similarity: 72.072      percent identity: 46.396

alignment_block:
US-09-581-241-4 x A1486799 ..

Align seq 1/1 to: A1486799 from: 1 to: 667

280 GlnAspTyrLysCysSerSerValIleLeuValProThrLeuPheAlaI 296
      |||||
1 CAAATTCAGAGGACCATAGGCGCATTTGTGCGACCTTATGTTTGGC 50
      ::::::::::::::|
296 eLeuAsnArgSerGluLeuLeuAspLysTyrAspLeuSerAsnLeuValG 313
      ::::::::::::::|
51 TATTGCTAAGAGTCCTATGCTGATATTAATTAATTCATCGGTAAAGA 100
      |||||
313 IuIleAlaSerGlyAlaProLeuSerLysGluIleGlyAlaVal 329
      ::::::::::::::|
101 CCGTTATGTCGGGGCTGCACCATTAAGAAAGAACTTGAAGACACGT 150
      ::::::::::::::|
330 AlArgArgPheAsnLeuProGlyValArgGlnGlyTyrGlyLeuThrG 346
      ::::::::::::::|
151 CGAGCCAAATTTCTTAATGCTAAACTTGTCAGAGTTACGAGATGACAG 200
      |||||
346 uThrThrSerAlaIleIleIle.....ThProGluGlyA 358
      ::::::::::::::|
201 AGCCGACGACAGTGTGCTATGCTTGGCATTTGCTAAAGAACCTTCG 250
      |||||
358 sPAspLysProGlyAlaSerGlyLysValValProLeuPheLysAlaLys 374
      ::::::::::::::|
251 AATTTAAATGTGACGCAATGCGACCTGTTGTAGAAATGCTCAGATGAAA 300
      |||||
375 ValIleAspLeuAspThrLysLysThrLeuGlyProAsnArgArgLys 391
      ::::::::::::::|
301 ATGTGATTCCTGATCTAGTGCACAAATCTCTCCCTAGAACCAATCTGAGA 350
      |||||
391 uValLysValLysGlyProMetLeuMetLysGlyTyrValAspAsnProG 408
      ::::::::::::::|
351 GATTTGTATAGAGGCGCATCAAAATTTAGAAAGTTACTTAATGATCCAG 400
      |||||
408 IuAlaThrArgGluIleIleAspGluGluGlyThrLeuHisThrLysAsp 424
      ::::::::::::::|
401 AGGCCACTACGCGCAACATAGAACAAAGAGGCTTACTATACGGCGCAT 450
      |||||
425 IieGlyTyrTyrAspGluGluLysHisPhePheIleValAspArgLeu 441
      |||||
451 ATTGTTATATTCATCATCATGATGAGCTTTTCATGTGATGCTTTAA 500
      |||||
441 sSerLeuIleLysTyrLysGlyTyrGlnValProProAlaGluLeuGln 458
      ::::::::::::::|
501 GGAATTCATTAATACAAAGATTTCAAGTGGCTCGCTGCAACTGGAAG 550
      |||||
458 erValLeuLeuGlnHisProAsnIlePheAspAlaGlyValAlaGlyVal 474
      ::::::::::::::|
551 CCTTCTCTCATATCATCCCAATATTTCAGATGCTGCTTGTTCATAG 600
      |||||
475 ProAspProIleAlaGlyLeuProGlyAlaValValLeuLysLys 491
      |||||
601 AAAGACGACCAAGCAGAGAGTCTCTGCGCTTTGTTGTACATCAAA 650
      |||||
491 sGlyLysSerMetThr 496
      ::::::::::::::|

```

```

651 TGGCTCCACAAATTACT 666
seq_name: gb_est2:BG414540
seq_documentation_block:
  LOCUS      BG414540      876 bp      mRNA
  DEFINITION  HVSMEK0002N03f Hordeum vulgare testa/pericarp EST library
  HYCDNA0013 (normal) Hordeum vulgare cDNA clone HVSMEK0002N03f, mRNA
  sequence.
ACCESSION   BG414540
VERSION     BG414540.1  GI:13320091
KEYWORDS    EST.
SOURCE      barley.
ORGANISM    Hordeum vulgare
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
             ; Triticeae; Hordeum.
REFERENCE   1 (bases 1 to 876)
AUTHORS    Wing,R., Close,T.J., Kleinbofs,A., Wise,R., Begum,D., Frisch,D., Yu
            ,Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo
            ,T., Saski,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and
            Wood,T.
TITLE       Development of a genetically and physically anchored EST resource
            for barley genomics
JOURNAL     Unpublished (2000)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: AATTAACCTCCTCACTAAGGG
            High quality sequence stop: 847.
            Location/Qualifiers
                source          1..876
                                /organism="Hordeum vulgare"
                                /cultivar="Morex"
                                /db_xref="taxon:4513"
                                /clone="HVSMEK0002N03f"
                                /clone_id="Hordeum vulgare testa/pericarp EST library
                                HVCDA0013 (normal)"
                                /HYCDNA0013 (normal)"
                                /tissue_type="testa/pericarp"
                                /lab_host="TJC121"
                                /note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI. For
                                more details on library preparation and sequence analysis
                                see http://www.genome.clemson.edu/projects/barley/ to
                                order a clone see http://www.genome.clemson.edu/orders"
BASE COUNT  248 a      187 c      209 g      232 t
ORIGIN
alignment_scores:
  Quality: 505.50      Length: 215
  Ratio: 3.082      Gaps: 4
  Percent Similarity: 76.279      Percent Identity: 49.767
alignment_block:
  US-09-581-241-4 x BG414540 ..
Align seg 1/1 to: BG414540 from: 1 to: 876
323 LysGluILlGLyGLuLaValAlaArgArGpHeAsnLeuProGLyValAr 339
      ::::::::::::::::::::::::::::::::::::::::::::
6 AGGAGCGTCATGAGGTGGGTGGCCAGAACTTCCCGACGCCGAGATCGT 55
339 gGlnGLyTYrGLyLeuThrGLuThrThrSerAlaIleIleIleThrProg 356
      |||||
56 TCAGGGCTATGATATGACTGAGACTGTGGG.....ATCATATCGCTTG 99
356 LucGLyspAspLysPro.....GLyAlaSerGLyLysVal 367
      ||
100 AATTAACCGAGAAAGGACAGGCTCGTCAATTGGGTCAACCGAAGCGTT 149

```

```

368 ValProLeuPheLysAlaLysValIleAspLeuAspThrLysThrIle 384
      |||
150 GTTGTGGAGTTGAGCAAAAAGTATGACGTAGAAAACGCAAAAACATCT 199
384 uGLyProAsnArGArgGLyGLuValLysValLysGLyProMetLeuMetL 401
      |||
200 ACCGCCAAGTCACTAGAGAAATCTATCTACGAGCCACACATATGCG 249
401 ySGLyTYrValAspAspProGLuAlaThrArgGluIleIleAspGLuGLu 417
      ::::::::::
250 AAGGCTATTTCACAAATGATAGAGCTACT...GAATTACAAATCGAGCA 296
418 GLYrLeuHisThrGLyAspIleGLYTYrTYrAspGLuLubLysHisph 434
297 GGATGGTTGCTACTGGTATCTTGATCTTGATCTTGATCAAGAGCCAGCT 346
434 ePheIleValAspArgLeuLysSerLeuIleLysTYrLysGLYTYrGLuV 451
      ::::::::::
347 TTTGTGTCGTGATAGACTAAAGAGCTGATTAGTACAAAGTTCCAGA 396
451 aLProProAlaGLuLeuGLuSerValLeuLeuGlnHisProAsnIlePhe 467
397 TTGCACCTCGCTGAGCTGAGATTAATCTTCTATCTCAATCAGATTCCTA 446
468 AspAlaGLyValAlaGLyValAlProAspProIleAlaGLyLubLeuProGL 484
447 GACGCCGTCGTATATCCGCTTCCTGATGATGAGCTGGGAGATTCCTAT 496
484 yAlaValAlaValLeuLysLysGLyLysSerMetThrGLuLysGLuValM 501
497 TGCCATATGGTGTAGAGTCACTGCCAGTTCATTAACGGAAGTGACGCTCC 546
501 eTAspTYrValAlaSerGLuValSerAsnAlaLysArgLeuArgGLyGLy 517
547 AAAAATTATGCGCAACAGGTGACGACTACTACAAACAAAGCTGAGG...CGG 593
518 ValArGpHeValAspGLuValAlProLysGLyLeuThrGLyLysIle 532
594 GTTACCTTCGTGAGAGCGTCCCAAAATCAAGCCGACGCAAGATC 638
seq_name: gb_est1:AW691106
seq_documentation_block:
  LOCUS      AW691106      656 bp      mRNA
  DEFINITION  NF04C08ST1F1000 Developing stem Medicago truncatula cDNA clone
  NE04C08ST 5', mRNA sequence.
ACCESSION   AW691106
VERSION     AW691106.1  GI:7565842
KEYWORDS    EST.
SOURCE      barrel medic.
ORGANISM    Medicago truncatula
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales;
             Medicago.
REFERENCE   1 (bases 1 to 656)
AUTHORS    He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
            ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
            ,R.A.
TITLE       Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula stem library
JOURNAL     Unpublished (2000)
COMMENT     Contact: Dixon RA
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7302
            Fax: 580 221 7380
            Email: radixon@noble.org
            Insert length: 656      Std Error: 0.00
            Plate: 041      row: C      column: 08
            Seq primer: TCACACAGCAACACGCTATGAC.

```

```

FEATURES
  source
    Location/Qualifiers
      1. 656
        /organism="Medicago truncatula"
        /db_xref="taxon:3880"
        /clone="MF041C08ST"
        /clone_lib="Developing stem"
        /tissue-type="stem"
        /dev_stage="Pooled developmental"
        /note="Vector: Lambda Zap; Contains a mixture of
        internal stem segments"
BASE COUNT      223 a      121 c      155 g      157 t
ORIGIN
alignment_scores:
  Quality: 504.50      Length: 219
  Ratio: 3.193      Gaps: 2
Percent Similarity: 72.146      Percent Identity: 46.575
alignment_block:
  US-09-581-241-4 x AW691106 ..
Align seg 1/1 to: AW691106 from: 1 to: 656
317 GlyGlyAlaProLeuSerLysGluIleGlyGluAlaValAlaArgArgph 333
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 GGAGCTGCTCCCATGGAATGGAACTGTAACACGCTGTTAAAGATAGTT 52
333 eAsnLeuProGlyValArgGlnGlyTYrGlyLeuThrGlu..... 346
53 GCCACGAACTGACTTGGACAGGATATGGATATGACAGAGGACGACACAC 102
347 ..ThrThrSerAlaIleIleThrProGluGlyAspAspLysProGly 362
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
103 TTTCATATTAGCTTGGCATTTGCCAAGAAAGAACCATTCAGACAAACCTGGC 152
363 AlaSerGlyLysValAlaProLeuPheLysAlaLysValIleAspLeuAs 379
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
153 GCATGTGGCAGTGTGAAGAAACGCTGAGTGAATAAGTAGTGACACAGA 202
379 pThrLysLysThrLeuGlyProAsnArgArgGlyGluValCysValLysG 396
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
203 AACTGTGCTTCCCTTCCAGAAAGAAAGCTGGTGAATTTGTTAGAG 252
396 LysProMetLysMetLysGlyTYrValAspAsnProGluAlaThrArgGlu 412
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
253 GCACCAAGCTTATGAAGAGATACCTAATAGACCTGAGCGACAAAGAGA 302
413 IleIleAspGluGlyTYrLeuPheLysThrGlyAspIleGlyTYrAs 429
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
303 ACTATAGACAAAGAGAGATGCTACACACAGGTGACATTGGTTAATTGA 352
429 pGluGluLysHisPhePheIleValAspArgLeuLysSerLeuIleLysT 446
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
353 CGACGACGAGAGAACTTTCATTTGTTGATGATTAAAGAAATTGATTAAAT 402
446 LysLysGlyTYrGluValProProAlaGluLeuGluSerValLeuLeuGln 462
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
403 ACAAGAGATACCAAGTAGCTGCTGAGCTCGAAGCATTTGATTGATGCT 452
463 HisProAsnIlePheAspAlaGlyValAlaGlyValProAspProIleAl 479
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
453 CATTCACAAACATTTCTGATGCTGCTGCTGATCCATTTGAAAGATGAAGCTGC 502
479 aglyGluLeuProGlyAlaValAlaValLeuLysLysGlyLysSerMetr 496
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
503 TGGAGAGAGTCCAGTGTGTTTGTGTAAGATCAAAATGTTCAAAGATCA 552
496 hrgLysGluValMetAspTYrValAlaSerGlnValSerAsnAlaLys 512
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
553 GCGAGGATGAATCAAGCAATACATATCCGACAGAGTTGATTTTACAAG 602
513 ArgLeuArgLysLysValAlaArgPheValAspGluValProLysGlyLeuTh 529

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
603 AGAATA...AACAGATTACTTACAGAAACAACTTCTTAAGCGCGCTC 649
529 rGlyLys 531
|||||:
650 AGGAAA 656
seq_name: gb_est1:A1485586
seq_documentation block:
LOCUS      A1485586      625 bp      mRNA      EST      18-MAY-2001
DEFINITION EST243907 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
CLEVED7A3, mRNA sequence.
ACCESSION  A1485586
VERSION    A1485586.1 GI:4380957
KEYWORDS   EST
SOURCE     tomato.
ORGANISM   Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 625)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujil,C.Y., Bowman
,C.L., Nieman,M., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
,S.D. and Giovannoni,J.
Generation of ESTs from tomato carpel tissue
Unpublished (1999)
CONTACT: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES
  source
    Location/Qualifiers
      1. 625
        /organism="Lycopersicon esculentum"
        /cultivar="7A496"
        /db_xref="taxon:4081"
        /clone="CLEVED7A3"
        /clone_lib="tomato ovary, TAMU"
        /tissue-type="carpel"
        /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
        /lab_host="XL1-Blue MRF"
        /note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:
        XhoI; cLED - Tomato Carpel EST Library. OligodT-primed and
        directionally cloned cDNA in vector Lambda Zap II with 5'
        and 3' ends located at the EcoRI and XhoI sites,
        respectively."
BASE COUNT      197 a      112 c      146 g      169 t      1 others
ORIGIN
alignment_scores:
  Quality: 482.50      Length: 200
  Ratio: 3.260      Gaps: 1
Percent Similarity: 74.000      Percent Identity: 48.500
alignment_block:
  US-09-581-241-4 x A1485586 ..
Align seg 1/1 to: A1485586 from: 1 to: 625
314 IleAlaSerGlyAlaProLeuSerLysGluIleGlyGluAlaValAl 330
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
17 GTATATGCTGGGGCTGCACCATTAAGCAAAAGAACTTGAACACACTGTTGC 66
330 aArgArgPheAsnLeuProGlyValArgGlnGlyTYrGlyLeuThrGluT 347
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
67 AGCCAAATTTCTAATGCTAAACTGTGTCAGAGTTACGAGATGACAGAG 116
347 hThrThrSerAlaIleIleIle.....ThrProGluGlyAsp 358
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```


117 CCGGACACAGTGTGGCTATGCTTGCGATTTGCTAAGAACCCCTCGAA 166
 359 AsplysProglYAlaSerGlyLysValProleuphelsAlaLysVa 375
 167 ATTAATCTGGAGCATGTGGACTGTGTAGAAATGCTCAGATGAAAT 216
 375 lLeaspleuasprhrlYsLysThrLeuGlyProasnarYargLysLuv 392
 217 TGTGGATCCTGTACTGGCAATCTCTCCAGAACCAATCTGGAGAGA 266
 392 aLysValLysGlyPrometleuMetLysGlyTYrValAspAsnProglu 408
 267 TTTGTAATAGAGCGCATCAATTAAGAAGTTACTTAATGATCCAGAG 316
 409 AlathrArgLuiIlelLeaspGluGlyTYrPleuHisThrGlyAspI 425
 317 GCCACTACGGGACAAATAGACAAAGAGGTGCTACATACGGGGATAT 366
 425 eGlyTYrTYrAspGluGlyLysHisPhePheIleValAspArgLeuLys 442
 367 TGTGTAATATTTGATCATGATGATGATGATGATGATGATGATGATG 416
 442 eLleuIleLysTYrLysGlyTYrGlnValProProAlaGluLeuGluSer 458
 417 AATTGATAAATACAAAGGATTTCAAGTGGCTCTGCTGAATCTGAAGCC 466
 459 ValLeuLeuGlnHisProAsnIlePheAspAlaGlyAlaLysValPr 475
 467 CTCTCTCTCAATCATCCCAATTTTCAGATGCTGCTGTTGTTCAATGA 516
 475 oAspProIleAlaGlyLysLysProglYAlaValAlaValLeuLysLysG 492
 517 AGACGAGCAAGCAAGAGAGTCTCTGCTGCTGCTGCTGCTGCTGCTG 566
 492 lYlSerMetThrGluLysGluValMetAspTYrValAlaSerGlnVal 508
 567 GCTCCCAATCTACTGAAATGCAAGATCATCAATCATCAATCAATCAAG 616
 seq_name: gb_est1.BE454671
 seq_documentation_block:
 LOCUS BE454671 686 bp mRNA EST 02-MAR-2001
 DEFINITION HVSMEH0094P16f Hordeum vulgare 5-45 DAP spike EST library
 HVGDM0009 (5 to 45 DAP) Hordeum vulgare cDNA clone HVSMEH0094P16f,
 mRNA sequence.
 ACCESSION BE454671
 VERSION BE454671.2 GI:13189489
 KEYWORDS
 SOURCE barley.
 ORGANISM Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Hordeum.
 1 (bases 1 to 686)
 Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
 'Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo
 'T., Saski,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and
 Wood,T.
 Development of a genetically and physically anchored EST resource
 for barley genomics
 Unpublished (2000)
 On Jul 26, 2000 this sequence version replaced gi:9463485.
 Contact: Wing RA
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: AATTAACTCATCTAAAGG
 High quality sequence stop: 653.
 Location/Qualifiers
 source 1..686

/organism="Hordeum vulgare"
 /cultivar="Morex"
 /db_xref="taxon:4513"
 /clone="HVSMEH0094P16f"
 /clone_lib="Hordeum vulgare 5-45 DAP spike EST library
 HVGDM0009 (5 to 45 DAP)"
 /tissue_type="5-45 DAP spike"
 /lab_host="SOLR"
 /note="Vector: lambdaZAP, site_1: EcoRI, site_2: XhoI. For
 more details on library preparation and sequence analysis
 see <http://www.genome.clemson.edu/projects/barley/> To
 order a clone see <http://www.genome.clemson.edu/orders>"
 BASE COUNT 156 a 212 c 203 g 115 t
 ORIGIN
 alignment_scores:
 Quality: 482.00 Length: 223
 Ratio: 2.957 Gaps: 2
 Percent Similarity: 73.094 Percent Identity: 43.049
 alignment_block:
 US-09-581-241-4 x BE454671 ..
 Align seg 1/1 to: BE454671 from: 1 to: 686
 314. lLealaserGlyAlaProleuSerLysGluIleGlyGlnAlaValAl 330
 17 GTCATGTCCGGCGCGCCCATGCGCAAGAGCTGACGACGCTTCAT 66
 330 aArgArPheasneLupProglYAlaArgGlnGlyTYrGlyLeuThrLut 347
 67 GGCCAAGATCCCAAGCGCGTCTCGGAGAGGGTACGGGATGACGAGG 116
 347 hrThrSerAlaIlellelelThrProGluGlyAspAspLysPro..... 361
 117 CCGGCGCGCTGCGCATGTGCTGCGCTGCGCTGCGCAAGAGGACCTTCAG 166
 362GlyAlaserGlyLysValAlaValProleuphelsAlaLysVa 375
 167 GTCAAGTCCGGGCTGCGCGCACGCTGCTGCGCAAGCGCGGCTCAAGAT 216
 375 lLeaspleuasprhrlYsLysThrLeuGlyProasnarYargLysLuv 392
 217 CGTGACCCCGACACCGCGCTGCTGCGCGGAGACACGCGCGGAGAGA 266
 392 aLysValLysGlyPrometleuMetLysGlyTYrValAlaAspAsnProglu 408
 267 TGTGATCCGCGGAGGACGATCAAGAGGTACCTTAACGACCCAGAA 316
 409 AlathrArgLuiIlelLeaspGluGlyTYrPleuHisThrGlyAspI 425
 317 TCACCAAGAACACCATGACAGATGCTGCTGCGACCGGAGACAT 366
 425 eGlyTYrTYrAspGluGlyLysHisPhePheIleValAspArgLeuLys 442
 367 CGGACATCTGCGGAGAGTCCGATGCTGCTGCTGCTGCTGCTGCTG 416
 442 eLleuIleLysTYrLysGlyTYrGlnValProProAlaGluLeuGluSer 458
 417 AGATCATCAAGTACAAAGGCTTCCAGGTGCGCGCGGACGACTCGAGGCC 466
 459 ValLeuLeuGlnHisProAsnIlePheAspAlaGlyAlaLysValPr 475
 467 CTCTCATACCCACCGGAGATCAAGACCGCGCGCTGATGCTGA 516
 475 oAspProIleAlaGlyLysLysProglYAlaValAlaValLeuLysLysG 492
 517 GGACGATCTTCTGCGGAGAGTCCGATGCTGCTGCTGCTGCTGCTG 566
 492 lYlSerMetThrGluLysGluValMetAspTYrValAlaSerGlnVal 508
 567 GCTTGAATTCACGAGATGATCAAAAATTTGTCGGAAGAGGTT 616

[illegible]

VERSION BF064865.2 GI:13266207
 KEYWORDS EST.
 SOURCE barley.
 ORGANISM Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Hordeum.
 REFERENCE 1 (bases 1 to 806)
 AUTHORS Wang, R., Close, T.J., Kleinholz, A., Wise, R., Begum, D., Frisch, D., Xu
 , Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo
 , T., Sasaki, C., Schwartzbeck, J., Simmons, J., Choi, D.W., Math, D. and
 Wood, T.
 TITLE Development of a genetically and physically anchored EST resource
 for barley genomics
 JOURNAL Unpublished (2000)
 COMMENT On Oct 17, 2000 this sequence version replaced gi:10841504.
 CONTACT: Wang RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: twang@clemson.edu
 Seq primer: AATTAACCTCCTCAAGG
 High quality sequence stop: 759.
 FEATURES
 source
 location/Qualifiers
 1..806
 /organism="Hordeum vulgare"
 /cultivar="C116151 (M1a6)"
 /db_xref="taxon:4513"
 /clone="HV_CEP0021D12f"
 /clone_lib="Hordeum vulgare seedling green leaf EST
 library HVCNDA0005 (Erysiphe infected & control)"
 /tissue_type="seedling green leaf"
 /lab_host="SOLR"
 /note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI; For
 more details on library preparation and sequence analysis
 see http://www.genome.clemson.edu/projects/barley/ To
 order a clone see http://www.genome.clemson.edu/orders"
 BASE COUNT 185 a 201 c 239 g 181 t
 ORIGIN
 alignment_scores:
 Quality: 479.00 Length: 211
 Ratio: 3.051 Gaps: 3
 Percent Similarity: 74.408 Percent Identity: 45.972
 alignment_block:
 US-09-581-241-4 x BF064865 ..
 Align seg 1/1 to: BF064865 from: 1 to: 806
 340 GlnGlyTyrGlyLeuThrGluThrThrSerAlaIleIleIleThrProG1 356
 ||||||||||||||||||||||||||||||||||||||||
 3 GAGGGGTAGCGGCTGACCGAGACGAGCGCGCGGCTCCACGAGCTC 52
 356 uGlyAspAsp.....LysProGlyAlaSerGlyLysValValProLeup 371
 ||||||||||||||||||||||||||||||||||||||||
 53 GGGCGGAGAGAGCGCGGCTGACGAGCGCGGCTCTGTCGCCACACA 102
 371 heLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyProAsn 387
 ||||||||||||||||||||||||||||||||||||||||
 103 CCGAGGCCAAGATGCTGACCGCGAGACGAGCGCGCTCGCGGAGAAC 152
 388 ArgArgGlyGluValLysValLysGlyProMetLeuMetLysGlyTyrVa 404
 ||| ||||||||: ||||||||| |||||||||
 153 CGACGCGGAGCTCTGATCAGGAGGACCTACGTCATGAAGATACCTT 202
 404 LAspAspProGluAlaThrArgGluIleIleAspGluGluGlyTyrLeuH 421
 ||| |||||||||: ||||||||| |||||||||
 203 CAAGAACACAGAGGCAACACATGCAAGGTCGCCCGCGGATGCTCA 252

421 IsThrGlyAspIleGlyTyrTyrAspGluGluLysHisPheIleVal 437
 |||||||||: ||| |||||||||: |||||||||
 253 AGACCGGTGACCTATGCTATACATACGAGAGATGGTATCTCTGTCGTC 302
 438 AspArgLeuLysSerLeuIleLysTyrGlyGlyTyrGluValProAl 454
 |||||||||: |||||||||: |||||||||
 303 GACGCTCGAAGAGATTGATCAATATACAAAGGATATCAGTGCCTCAGC 352
 454 agLLeuGluSerValLeuLeuGluGlnHisProAsnIlePheAspAlaGly 471
 |||||||||: |||||||||: |||||||||
 353 AGACTTTGCAACTCTCTGCTGACCATCCAGAGCTTTCGAGTGCCTG 402
 471 AlaAlaGlyValProAspProIleAlaGlyGluLeuProGlyAlaVal 487
 || ||||| ||||| ||||| ||||| |||||
 403 TTATTCCTTCCTCCACAGACGAGGTCGTCAGTCCGAGATGCGCTATGTC 452
 488 ValLeuLysGlyGlyLysSerMetThrGluLysGluValMetAspTyrVa 504
 || |||||||||: |||||||||: |||||||||
 453 GTGAGGAAGAAAGGAGCAATTGTGACGCGAGAGGTGATGAGATTGT 502
 504 LAlaSerGluValSerAsnAlaLysArgLeuArgGlyGlyValAlaArgPhe 521
 |||||: ||||| ||||| ||||| ||||| |||||
 503 GCGCAACAGGTAGACCTTCAAGAAAGGTAGG...AAGTGCGCTTCG 549
 521 AlaSpGluValProLysGlyLeuThrGlyLysIleAspGlyLysAlaIle 537
 || |||||||||: |||||||||: |||||||||
 550 TGACAGACATCCCAAGATGCGTGTGCAAGATACGAGAGAG..... 593
 538 ArgGluIleLeuLysLysProValAlaLysMet 548
 |||||||||: |||||||||: |||||||||
 594 ...GATCTTATCAAGCTCGCAGCTCCACAGCTG 623
 seq_name: gb_estc:BE820931
 seq_documentation_block:
 LOCUS BE820931 772 bp mRNA 24-MAY-2001
 DEFINITION GW700013A20B2 Gm-r1070 Glycine max cDNA clone Gm-r1070-5044 3',
 mRNA sequence.
 ACCESSION BE820931
 VERSION BE820931.1 GI:10253165
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseolae;
 Glycine.
 REFERENCE 1 (bases 1 to 772)
 AUTHORS Vokkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,
 Expelling, J., Raph, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.
 TITLE A Functional Genomics Program for Soybean (NSF 9872565)
 JOURNAL Unpublished (1999)
 COMMENT Other-ESTs: AM458737 corresponding to Gm-cl0106-6629 (5')
 Contact: Vokkin, L.O., PI, A Functional Genomics Program for
 Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional
 Genomics
 University of Illinois
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582
 Email: l-vokkin@uiuc.edu
 This clone is available through: Genome Systems, Inc. 4633 World
 Parkway Circle St. Louis, Missouri 63134. For further information
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or info@genome
 systems.com web site: www.genomesystems.com
 Seq primer: 5'-TTTATTTTATTTTATTTT(A/C/G)-3'.
 FEATURES
 source
 location/Qualifiers
 1..772
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="Gm-r1070-5044"

/clone_lib="Gm-r1070"
 /note="The library Gm-r1070 is a sequence-driven, reracked set of 9,216 clones selected from cDNA libraries from various tissues and stages of development of soybean that represent 2,639 sequences from immature cotyledons, 1,770 from immature seed coats, 3,938 from flowers, and 869 from young pods. The 5' ESTs of the source clones from the different libraries was used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1070. The cDNA clones of the reracked Gm-r1070 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the Laboratory of Ernest Reizel, Center for Computational Genomics and Bioinformatics, University of Minnesota, <http://www.cbc.umn.edu/Research/Projects/Soybean/Index.html>. Reracking was performed by Genome Systems, St. Louis, <http://www.genomesystems.com>, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, <http://www.life.uiuc.edu/biotech/keck.html>. Note: The corresponding 5' EST from each clone in the Gm-r1070 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Genome Systems clone ID of the original cDNA library that is also listed under 'OTHER EST'."

BASE COUNT 195 a 190 c 143 g 214 t 30 others
 ORIGIN

alignment_scores: Quality: 477.50 Length: 243
 Ratio: 2.776 Gaps: 5
 Percent Similarity: 70.782 Percent Identity: 43.210

alignment_block:
 US-09-581-241-4 x BE820931/rev ..

Align seg 1/1 to reverse of: BE820931 from: 1 to: 772

```

307 AspleuSeranLeuValGluilelaSerclyglValaProleuSerly 323
    ||| :||||| |||:
747 GATNNNNCGTCGCGGANNNGATGTTGCCNNNGATCTCCGCGAGAGAN 698
    ||| :||||| |||:
323 sgluileglValaValaIaArgArpheaSleuProglYvalArg 340
    ||| :||||| |||:
697 ANNNNNAGGAGGAGCATTAAGCTCAAGCTCCCAAGCTCCTCATATGC 648
    ||| :||||| |||:
340 lnglyTyrlyleuThrcluthrThrSeralallele...IleThrPro 355
    ||| :||||| |||:
647 AGGATATCGGTTAAATGAG.....TCTGCAGTTACCCGAAACCATCCA 604
    ||| :||||| |||:
356 glUGlyAspAspLysProglYalaserGlyLysValaValProleuPhe 372
    ||| :||||| |||:
603 GAGGAGCAATATCAGTGGGAGCAACAGGTAACTGATACCAATATAGA 554
    ||| :||||| |||:
372 salalysValalleaspleuaspThrlyLysThrleuGlyProasnar 389
    ||| :||||| |||:
553 AGCCAAATATGTGAACCAAGCAAGGAGCATGTTCCCGGTGAC 504
    ||| :||||| |||:
389 rglyglValaLysValaLysGlyProleuLewetLysGlyTyrVala 405
    ||| :||||| |||:
503 AAGGAGAACTCTGATCAGAGACCTTATGTATGAAGAGTTATTCGGT 454
    ||| :||||| |||:
406 AspProgluaIaThrArgGluileleaspgluGluYtrpleuIsth 422
    ||| :||||| |||:
453 GACCAAAACCACTTCAGCACTTTC...GTGATGCGGTGTTAAGGAC 407
    ||| :||||| |||:
422 rglYasplleGlyTyrAspGluLysShsPhePheIleValaAsp 439
    ||| :||||| |||:
406 TGGGAGACCTCTGATATTTGATAGCAAAAGTTCTTATGTGTAGATA 357
    ||| :||||| |||:
439 rglleuysSerleuileLysTyrlyglValaProleuIaGlu 455
  
```

```

356 GGTGAAAGAGTGTATTAATACAAAGGCTACAGGTGCTCCGCGAGAG 307
    ||| :||||| |||:
456 leuGluSerValleuLeuGlnHisProasnllePheaspAlaGlyVala 472
    ||| :||||| |||:
306 CTGAAAGAGTATCTTCATCTCATTCAGAGATAATGATGCTGCAGTAT 257
    ||| :||||| |||:
472 aglyValProasproIlealaglyGluLeuProglYalValaVala 489
    ||| :||||| |||:
256 TCATATACCTCATGATAGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 207
    ||| :||||| |||:
489 euLysGlyLysSerMetThrGluLysGluValaMetasPtyrVala 505
    ||| :||||| |||:
206 GACAACCTCAAGATTCCTTGGCGGAGCAAGATGATGATTTGTTGCC 157
    ||| :||||| |||:
506 SerGluValaSeranAlaLysArgLeuArglyGlyValaArgPheVala 522
    ||| :||||| |||:
156 AAGAGGTTTCACCGTACACAAATAAGG...CGTATGATTTGTTAA 110
    ||| :||||| |||:
522 pglValaProLysGlyLeuThrGlyLysIleaspglyLysAlaIleArg 539
    ||| :||||| |||:
109 TTCCATACCCAGAAATGCTGCGGAGAAAGAT.....CTGAGA 72
    ||| :||||| |||:
539 luileLysLysProValaLysMet 548
    ||| :||||| |||:
71 AGACTTAACAACTGCGCTCTCTAGGTTG 43
    ||| :||||| |||:
seq_name: gb_est1:AV558307
  
```

seq_documentation_block:

LOCUS AV558307 604 bp mRNA EST 07-SEP-2000
 DEFINITION AV558307 Arabidopsis thaliana green siliques Columbia Arabidopsis
 thaliana cDNA clone S094e05F 3', mRNA sequence.

ACCESSION AV558307
 VERSION AV558307.1 GI:8729733

KEYWORDS EST.
 SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 604)
 Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.

A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries

JOURNAL DNA Res. 7, 175-180 (2000)
 MEDLINE 20363093

COMMENT Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES
 source location/Qualifiers
 1..604

/organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="S094e05F"
 /clone_lib="Arabidopsis thaliana green siliques Columbia"
 /clone_type="green siliques"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI"

BASE COUNT 165 a 129 c 148 g 162 t
 ORIGIN

alignment_scores: Quality: 474.50 Length: 203
 Ratio: 3.142 Gaps: 2
 Percent Similarity: 74.384 Percent Identity: 47.291

alignment_block:

US-09-581-241-4 X AV558307

Align seg 1/1 to: AV558307 from: 1 to: 604

[illegible]


```

223 salAaTgAspProIleTyrGlyAsnGlnValSerProGlyThrAlaIleL 240
284 CTCGAAGAAGTGGGCTAAGCTGCTGAGGCTACCCGATGAGCTGTGACAC 333
240 eutThValValProPheHisHis.....GlyPheGlyMetPheThr 253
334 TAGCAGTGTACTCTCTGTTCCAGCTGCGCCGCTGTCTGCGGCTG..... 377
254 ThrLeuGlyThrLeuThrCysGlyPheArgIleValMetLeuThrIlePhe 270
378 ...CTGAATATGTTATCCAAACATTCCTCGGTGTGTGTGATGCCAGATT 424
270 eAspGluGluThrPheLeuLysThrLeuGlnAspTyrLysCysSerSery 287
425 TCATCCACACCTGTTCTTACGTATGTAAGAATATCGGTAATCTAA 474
287 aIleLeuValProThrLeuPheAlaIleLeuAsnArgSerGlyLeuLeu 303
475 TCACACTGTACCTCCGCTGATGCTCTTCTGCCCCAACATCCCATGGTG 524
304 AspLysTyrAspLeuSerAsnLeuValGluIleAlaSerGlyValAlaPr 320
525 GACACTAGCATCTCTCTCGCTGATGAGCTGTCTGTGCGCGGCTCC 574
320 OLeuSerLysGluIleGlyAlaValAlaArgArgPheAsnLeuProG 337
575 GTTGAGCAAGAAGATCAGATCAGGTGCGGAGCAGCATGGGTAGCTT 624
337 lValAlaArgGlnGlyTyrGlyLeuThrGluThrThrSerAlaIleIle 353
625 TCATTTCGAGGATGAGTACGATGAGTGAACCAACCTCGTGTGCTGATG 674
354 ThrProGluGlyAspAspLysProGlyAlaSerGlyLysValValPro 370
675 CAGGACGAGTTGAGAACAAAGCGGACGCTGGGAAAGTGGCGCATGGG 724
370 uPheLysAlaLysValIleAspLeuAspThrLysLysThrLeuGlyPro 387
725 CCAGTGGGTGAAGGTGATGACCCGAGGAGCGGCGCTGCGCGACGA 774
387 sAlaArgGlyGlyValLysValLysGlyProMetLeuMetLysGlyTyr 403
775 ACCAGCGGGGTAGCTGCTTTAAGGCTCATGTATGAAGAAGATAT 824
404 ValAspAsnProGluAlaThrArgGluIleIleAspGluGluGlyTyr 420
825 GTGGCGAAGAGCAGGCC.....ATCGATGCTGACGCTGGCT 862
420 uHisThrGlyAspIleGlyTyrTyrAspGluGluLysHisPhePheIle 437
863 GCACACGGGCGACATTGCTACTACATGATGATCAGAGTTTTCATCG 912
437 aAlaAspArgLeuLysSerLeuIleLysTyrLysGlyTyrGlnValPro 453
913 TGCACCGGATCAAGAGCTGATCAAGAGGCTTCCAGTACCGCGG 962
454 AlaGluLeuGlu 457
963 GCCGAGCTCGAA 974

seq_name: gb_est2:BG445463

seq_documentation_block:
LOCUS BG445463 882 bp mRNA EST 15-MAR-2001
DEFINITION GA_Ea0028K14f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboresum cDNA clone GA_Ea0028K14f, mRNA sequence.
ACCESSION BG445463
VERSION BG445463.1 GI:13355211
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum;
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

REFERENCE 1 (bases 1 to 882)
AUTHORS Wing,R.A., Fritsch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE An integrated analysis of the genetics, development, and evolution
of the cotton fiber
JOURNAL Unpublished (2000)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCACTATAGG
High quality sequence stop: 733.
Location/Qualifiers
1..882
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0028K14f"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/Note="vector: pBK-CMV, Site_1: EcoRI, Site_2: XhoI"
BASE COUNT 250 a 173 c 232 g 226 t 1 others
ORIGIN

alignment_scores:
Quality: 552.00 Length: 277
Ratio: 2.816 Gaps: 4
Percent Similarity: 70.758 Percent Identity: 41.155

alignment_block:
US-09-581-241-6 x BG445463 ..

Align seg 1/1 to: BG445463 from: 1 to: 882
263 ArgIleValMetLeuThrLysPheAspGluGluThrPheLeuLysThrLe 279
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 CGAGTTTGTATCATGCGAAGATTGCGACATCGCGCTTGTGAGCTAAT 55
279 uGlnAspTyrLysCysSerSerValIleLeuValProThrLeuPheAlaI 296
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
56 ACAGAAATACAAAGTACCATGCTGCTCGATTGCGCCAGATGTTTGG 105
296 lLeuAsnArgSerGluLeuLeuAspLysTyrAspLeuSerAsnLeuVal 312
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 CCATCGCTAAGTCATCGGAAGTGAAGAAATTTAGCATTTGTCATCGGAGG 155
313 GluIleAlaSerGlyGlyAlaProLeuSerLysGluIleGlyGluAlaVa 329
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
156 ANGTGAAGTCCGCTCCGCTCGCTGGGTCAAGAGCTTGAAGATCTGT 205
329 lAlaArgArgPheAsnLeuProGlyValArg.....GlnGlyTyrGlyL 344
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 AAAAGTCAAGTTT.....CTGTGTCCAAAATTTGACAGAGGTTAAGAA 249
344 eutThrGluThrThrSerAlaIleIleIle.....ThrPro 355
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
250 TGAAGAAAGCTGAGCCAGATTCTTACCAATGTGTTGGATTTGCCAAGAA 299
356 GluGlyAspAspLysProGlyAlaSerGlyLysValValProLeuPheL 372
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
300 CCTTTGAATGAATCCGAACTTGTGGAGCTGTTGTAAGAACCGGA 349
372 salAlaLysValIleAspLeuAspThrLysLysThrLeuGluGlyProAsnArg 389
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



```

350 GATGAAATATGTCGACCCGAAACCGGTTTGTCTTCGCCGACAGACCAGG 399
389 rgllygluValCysValLysGlyProMetLeuMetLysGlyTyrValAsp 405
400 CCGGCGAGATTTCATTCGAGAGATCATCATGAAAGATACCTTAAT 449
406 AsnProGluValThrArgGluLeuLeuAspGluGluLysPheLeuHisThr 422
450 GACCTGAGGCGACTGCTAGACCATTCAGACAAGATGGCTGGTTACTATC 499
422 rglYAspIleGlyTyrTyrAspGluGluLysHisPheLeuValAspa 439
500 CGGAGACATTGCTTACATTCAGCGATGATGAACCTTCATCGTCGATC 549
439 rglLeuLysSerLeuIleLysTyrLysGlyTyrGluValProPheValGlu 455
550 GATTGAAAGAAATGATTAATAACAAAGGTTTCAGTTGCCCTGCCGAG 599
456 LeuGluSerValLeuLeuGluHisProAsnIlePheAspAlaGlyValAl 472
600 CTCGAGNAGCATGCTCATTCGTCACCCGAAATCATCATGCTGCTGCTG 649
472 aglyValProAspProIleAlaGlyGluLeuProGlyAlaValAlValL 489
650 CGCATATGAGGATGAGACAGCTGGGGAAGTACCTGTGCACTTGTGTGTA 699
489 euLysLysGlyLysSerMetThrGluLysGluValMetAspTyrValAla 505
700 GATCATGATTAATCTCATGATCATGATGATTAATCAAGCAATATATTTCG 749
506 SerGluValSerAsnAlaLysArgLeuArgLysIleValAlaArgPheVal 522
750 AAACAGGAGGTGTTCTATAGGAGATTAAGCCGGGGTTC...TTCAATTGG 796
522 pgluValProLysGlyLeuThrGlyLysIle 532
797 AGCCCTTCCTCAAGGCCCATCATCAGGAGGAATC 827

```

seq_name: gb_est1:BE036528

```

seq_documentation block: 1066 bp mRNA EST 07-JUN-2000
LOCUS BE036528
DEFINITION MP01A02 MP Mesembryanthemum crystallinum cDNA 5' similar to
coulmarate:coenzyme a ligase, mRNA sequence.
ACCESSION BE036528
VERSION BE036528.1 GI:8331537
KEYWORDS EST.
SOURCE common ice plant.
ORGANISM Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
Caryophyllales; Caryophyllaceae; Alzooceae; Mesembryanthemum.
1 (bases 1 to 1066)
Bohnert,H.O., Borchert,C., Brazille,S., Brooks,J., Eaton,M., Ferreira
,H., Kawasaki,S., McCollough,A., Michalowski,C.B., Palacio,C.,
Scara,G., Wheeler,M. and Zepeda,G.R.
Functional Genomics of Plant Stress Tolerance
Unpublished (2000)
Contact: Michalowski,C.B.
University of Arizona
Bio Sciences West room 513, Tucson, AZ 85721, USA
Tel: 520-621-7982
Fax: 520-621-1697
Email: cbm@u.arizona.edu
An open reading frame exists.
Insert Length: 1 Std Error: 0.00.

```

FEATURES

```

source
1..1066
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone_lib="MP"
/tissue_type="apical meristem and leaf primordia"
/dev_stage="6 weeks"

```

```

BASE COUNT 265 a 220 c 279 g 294 t 8 others
ORIGIN
alignment_scores:
    quality: 541.00    length: 295
    ratio: 2.678
    percent similarity: 68.475    percent identity: 40.339

```

alignment_block:

US-09-581-241-6 x BE036528 ..

Align seg 1/1 to: BE036528 from: 1 to: 1066

```

218 ValThrArgPhe.....SerHisAlaArgAspProIleTyrGlyAsnG1 232
4 ATTAGCGCTATATGAGTGCATATTAATATCCG..... 39
232 nValSerProGlyThrAlaIleLeuThrValValProPheHisHisGlyP 249
40 .....CCAAAGCGCGTCTCTCTCCCGCTTTCATGCTCT 76
249 heGlyMetPheThrThrLeuGlyTyrLeuThrCysGlyPheArgIleVal 265
77 TTGGTTCTTCATGCTGCTCAGAGCGGTGCTGCGCCGAACTAATAGGTT 126
266 MetLeuThrLysPheAspGluGluThrPheLeuLysThrLeuGluAspTyr 282
127 GTATGGCAAGGTTGATTCGTAAGATGTTGGAGGCTGTGGAGAGTA 176
282 rLysCysSerSerValIleLeuValProThrLeuPheAlaIleLeuAsna 299
177 TCGGATATACATACATCCGTTTCGCCGATGTTGGTGGCGATGACAA 226
299 rSerGluLeuLeuAspLysTyrAspLeuSerAsnLeuValGluIleAla 315
227 AATCGAGGCTTACAGATTAAGTACGATCTCAGCAGCGTGAAGTGTCT 276
316 SerGlyValAlaProLeuSerLysGluIleGlyGluAlaValAlaArgAr 332
277 TGTGTGTGCGCTCTCTCTGGGAGGAGGTTGCACACGCTTCGACTG 326
332 gPheAsnLeuProGlyValAlaArgGlnGlyTyrGlyLeuThrGluThrHis 349
327 CTTCGCCGAACTGAGATACACAGGATATAGGCTGACAGACAGCAG 376
349 eValAlaIleIleThr.....ProGluLysAspLysProGlyAla 363
377 GCCCGGCTGCAGAGAGCTTATAGGCTGAGGAAATCTGACGATATGGTCT 426
364 SerGlyLysValValProLeuPheLysAlaLysValIleAspLeuAspThr 380
427 GTTGGCGCGCTTACAGAGCTCTTGGAGCAAGATATGATTCGCTCAAC 476
380 rLysLysThrLeuGlyProAsnArgArgGlyGluValCysValLysGlyP 397
477 CGGAGAGCGCTTACACCTGCTGCAAGAGGAGGAGCTGTGGCTCGGGGC 526
397 rMetLeuMetLysGlyTyrValAlaAspAsnProGluValThrArgGluIle 413
527 CAACCATTAATGAAGAAAGCTATGTGGAGATTAAGATCCCATCTGCAACT 576
414 IleAspGluGluGlyTyrPheLeuHisThrGlyAspIleGlyTyrTyrAspG1 430
577 NTCCATTCANATGATGATGCTGCAAGACTGATGATCTTTGATTTTACATC 626
430 uGluLysHisPhePheIleValAspArgLeuLysSerLeuIleLysTyrL 447
627 TGATGCGTTCTTCATATGTTGATGATTAAGATTAAGATTAAGATTAAG 676
447 ysgLysTyrGluValProPheAlaGluLeuGluSerValLeuLeuGluHis 463

```

```

677 AGGCATATCATGNGTTCCCTCAGCGTGAAGTGAGGCGTTACTACATGCCAT 726
464 ProAn1LpheaspAlagLYValAlAgiLYValProAspProllealgl 480
|||||:||||| ||||| :||| |||||
727 CCAGATTATTCGTGATGCTGCTGGTGTGCCGTATCCAGATACCAAGCG 776
480 yGuLeuPrgGLyAlaValAlValleuLySlySGLY.LysSerMetThr 496
|||||:||||| ||| ||||| :||| |||
777 ACAAAATACCATTGCGATTGTGGTC..AGCAATNCAGAACCAACCATTTTC 823
497 GluLYsgLUValMeTasPYrValAlASerGLn 507
|||||:||||| ||||| :||| |||||
824 TGCTCTCACCTCATGATATGTAAGCTAACAGC 856

seq_name: gb_estl:AU214129

seq_documentation_block:
LOCUS AU214129 741 bp mRNA EST 17-JUL-2001
DEFINITION AU214129 unpublished oligo-capped cDNA library, stage L2
Accession CAenorhabditis elegans CDNA clone yX809f05 3', mRNA sequence.
AU214129
VERSION AU214129.1 GI:14852286
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea.
1 (bases 1 to 741)
Kohara,Y., Shlin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
A complementary view of the C.elegans genome
Unpublished (2001)
Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: yokohara@lab.nig.ac.jp.
Location/Qualifiers
1..741
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yx809f05"
/cclone_lib="unpublished oligo-capped cdna library, stage
L2"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L2"
BASE COUNT 216 a 158 c 125 g 241 t 1 others
ORIGIN

alignment_scores:
Quality: 513.00 Length: 223
Ratio: 3.090 Gaps: 5
Percent Similarity: 74.439 Percent Identity: 47.982

alignment_block:
US-09-581-241-6 x AU214129/rev ..

Align seg 1/1 to reverse of: AU214129 from: 1 to: 741

320 ProLeuSerLySgLUleGLyGUAlaValAlArGaRPhE...ASole 335
|||||:|||||:||||| |||:|||||:|||||
739 CCAGCTGGAAGAAGATTTGATGACAAGATTGAAGAACTATACAAATTT 690
335 uProGLyVALarGInGLyTYrGLYleuthrGUthrPrSerAlaleI 352
|||||:|||||:||||| |||:|||||:|||||
689 GAATATACATCCAACAAGATATGGAATGACGGAATGTTCAATAGCGTAGTC 640
352 leilethrProGLUGlyaspLySPro.GLYAlASerGLyLYsVAL 367

```

seq_name: gp_est2:C83857

seq_documentation_block: 664 bp mRNA EST 28-APR-1999

LOCUS C83857

DEFINITION C83857 Dictyostelium discoideum SS (H. Urushihara) Dictyostelium

ACCESSION C83857

VERSION C83857

KEYWORDS C83857.1 GI:2706789

SOURCE EST.

ORGANISM Dictyostelium discoideum.

ORGANISM Dictyostelium discoideum.

ORGANISM Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

REFERENCE 1 (bases 1 to 664)

AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M., Yoshino,R., Mitera,B.N., Pl.M., Saito,T., Takemoto,K., Yasukawa,H., Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.

TITLE The Dictyostelium developmental cDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development

JOURNAL DNA Res. 5 (6), 335-340 (1998)

MEDLINE 99156227

COMMENT Contact: Hideko Urushihara

Institute of Biological Sciences

University of Tsukuba

3-3-10 Ten-nodal, Tsukuba, Ibaraki 305, Japan

Email: d402hu@sakura.cc.tsukuba.ac.jp

PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
POLYA-No.

FEATURES
source location/Qualifiers
1..664 /organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SSA391"
/clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
/dev_stage="slug"

BASE COUNT 234 a 105 c 124 g 201 t
ORIGIN

alignment_scores:

Quality: 512.00 Length: 200
Ratio: 3.261 Gaps: 0
Percent Similarity: 78.500 Percent Identity: 47.500

alignment_block:

US-09-581-241-6 x C83857 ..

Align seg 1/1 to: C83857 from: 1 to: 664

```

340 GlnGlyTyrGlyLeuThrGluThrThrSerAlaIleIleIleThrProG1 356
|||||
2 CAAGTTATGTTGCTACCGAAGCTTACCGAGCTGTTTGTATTCACAG 51
|||||
356 uGlyAspAspLysProGlyAlaSerGlyLysValProLeuPheLysA 373
|||||
52 TGGATTAGTTAAATCTGTTCCGCTGCTATCTTATCCCAATCAATTCG 101
|||||
373 lAlysValIleAspLeuAspThrLysThrLysLeuGlyProAsnArg 389
|||||
102 CTAAATCATCTCTCCAGAGCTGCTAGAGATTAGGTAGGGAAG 151
|||||
390 GlyLysValCysValLysGlyProMetLeuMetLysGlyTyrValAsp 406
|||||
152 GGTGAATTCGTATTAAGGCCAATGTTATGTTGTTATATATACAA 201
|||||
406 nProGluAlaThrArgLuiIleIleAspGluGluGlyTTrPheHisThr 423
|||||
202 TGAAGAAGCAACCAATGATCATGATGATGATGATGATGATGATGAT 251
|||||
423 lYAspIleGlyTyrTyrAspGluGluLysHisPhePheIleValAsp 439
|||||
252 GTGATTATGTTACGTCGATGATGATGATGATGATGATGATGATGAT 301
|||||
440 LeuLysSerLeuIleLysTyrLysGlyTyrGluValProProAlaGlu 456
|||||
302 TCAGAGGATTCATCAATGATTAAGGTTTCCAGTACTCTCTGCGAATT 351
|||||
456 uGluSerValLeuLeuGlnHisProAsnIlePheAspAlaGlyValA 473
|||||
352 GGAGGCAATTACTATTATCTCATCCAAAGTTGACAGATCTGCTAGTA 401
|||||
473 lYValProAspProIleAlaGlyLysLeuProGlyAlaValValLeu 489
|||||
402 GTCTTCCAAAGGTGATATGGGTACAGTACCAAGAGTTTCGTTGTTAT 451
|||||
490 LysLysGlyLysSerMetThrGluLysGluValMetAspTyrValAla 506
|||||
452 AAACAAATGCAATCTTAACCGAAGAACTCTTACATTTGGGCCATCC 501
|||||
506 rGlnValSerAsnAlaLysArgLeuArgGlyGlyValArgPheValAsp 523
|||||
502 AAAAATTGCAAAATTATAACATTTCAAGAGGTGATCTTTTTCATAC 551
|||||
523 lValProLysGlyLeuThrGlyLysIleAspGlyLysAlaIleArgGlu 539
|||||
552 CAATTTCCTAAATCTGCAACGCGTAACCTATTACGTAATAAATCTTA 601
|||||

```

seq_name: gb_est2:BF492687

seq_documentation_block:

LOCUS BF492687 683 bp mRNA EST 19-APR-2001

DEFINITION AT01136.3prime AT Drosophila melanogaster adult testes POTB7
Drosophila melanogaster cDNA clone AT01136 3 similar to CG6178:
FBan0006178 'enzyme' located on: 3R 95E1-95E1; 04/07/2001, mRNA
sequence.

ACCESSION BF492687
VERSION BF492687
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 683)
Stapleton,M., Brokstein,P., Hong,L., Agbayan,A., Baxter,E., Berman,
B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V., Farfan,
D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P.,
Liao,G., Miranda,A., Misra,S., Mungall,C.J., Nunoo,J., Pacle,J.,
Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E.,
Celisner,S. and Rubin,G.M.
BDGP/HMT AT Drosophila EST Project
Unpublished (2000)
Other ESTs: AT01136.5prime
Contact: Stapleton, M.

TITLE JOURNAL
COMMENT
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: <http://www.fruitfly.org/EST>
hit genomic AE003746: arm:3R [19764572,19982668]
estimated-cyto:95D5-95F1: 04/07/2001
Plate: AT.11 row: C column: 12
High quality sequence stop: 631.

FEATURES

source

location/Qualifiers
1..683
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="AT01136"
/clone_lib="AT Drosophila melanogaster adult testes POTB7"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/note="Organ: ADULT testes. Vector: POTB7; Site:1. ECKR1:
Site:2: Xho1; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into POTB7. Plasmid cDNA library."

BASE COUNT 169 a 188 c 149 g 177 t
ORIGIN

alignment_scores:

Quality: 507.50 Length: 176
Ratio: 3.574 Gaps: 1
Percent Similarity: 80.682 Percent Identity: 57.955

alignment_block:

US-09-581-241-6 x BF492687/rev ..

Align seg 1/1 to reverse of: BF492687 from: 1 to: 683

```

373 lAlysValIleAspLeuAspThrLysThrLysLeuGlyProAsnArg 389
|||||
637 GCCAAGGTGATCGATCCGACACCGCAAGCTATTGCGCCACACGAG 588
|||||
389 gGlyGluValCysValLysGlyProMetLeuMetLysGlyTyrValAsp 406
|||||
587 CGCGAGCTTGTTTAAAGCGACGCAATCATGAAGGCTACATCGAG 538
|||||

```

```

406 snProgluAlaThrArgGluIleIleAspGluGluGlyTrpLeuHisThr 422
    :: :::::::::::::: ::::::::::::::
537 ATACGAAAGTCACGACGAGACCGGCATC...AAGGACGGTGGTGCATCT 491
423 GlyAspIleGlyTyrTrpAspGluGluGlyHisPhePheIleValAsp 439
    ||||| :::::::::::::: ::::::::::::::
490 GCGATATTGGCTACTATATGATGATTTTGAATTCCTTCATCCTGACCG 441
439 gLeuLysSerLeuIleLysTyrTrpGlyTyrGluValProProAlaGlu 456
    |::| :::::::::::::: ::::::::::::::
440 CATCAAGAGAGCTGATCAATACAGGATACAGGATCGCCGCCGCGACAG 391
456 euGluSerValLeuGluGlnHisProAsnIlePheAspAlaGlyAla 472
    :::::::::::::: ::::::::::::::
390 TTGAGGCTCTGCTGCTACCAACGACAGATTAAGATGCGCGGCTCAT 341
473 GlyValProAspProIleAlaGlyGluLeuProGlyAlaValValle 489
    ||| :::::::::::::: ::::::::::::::
340 GGAAGCCGACGACGAGAGCTGCGAGCTGCCGCTGCTGCTGCTCAA 291
489 uLysLysGlyLysSerMetThrGluLysGluValMetAspTyrValAl 506
    :: :::::::::::::: ::::::::::::::
290 ACAGGCTATATGTTCACTGACCGACGACGACGATTCAGTTGTCAAG 241
506 ergLysIleSerAsnAlaLysArgLeuArgGlyGlyValArgPheVal 522
    :::::::::::::: ::::::::::::::
240 ACAAGCGCTCGCCGCCGACGCTCTAAGGGGTCGCTGATCTTTGTGAC 191
523 GluValProLysGlyLeuThrGlyLysIleAspGlyLysAlaIleArg 539
    ||::| :::::::::::::: ::::::::::::::
190 GAATATTCACAAAGAACCCAGTGGCAAGATTCGCGTCGATTCGCGGA 141
539 uLleLeuLysLysProValAlaLysMet 548
    |::| ::::::::::::::
140 AATGCTTAAGAAAGCAAAATCCACAGTTG 113

seq_name: gb_est1.A1486799

seq_documentation_block:
LOCUS      A1486799      667 bp      mRNA      EST      18-MAY-2001
DEFINITION EST245121 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
            cLED11D3, mRNA sequence.
ACCESSION  A1486799
VERSION    A1486799.1 GI:4382170
KEYWORDS   EST.
SOURCE     tomato.
ORGANISM   Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; eusterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE  1 (bases 1 to 667)
            Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,
            Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
            ,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
            ,S.D. and Giovannoni,J.
            Generation of ESTs from tomato carpel tissue
            Unpublished (1999)
            Contact: CUGI
            Clemson University Genomics Institute
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html.

FEATURES             source
     source
         1..667
         /organism="Lycopersicon esculentum"
         /cultivar="7496"
         /db_xref="taxon:4081"
         /clone="cLED11D3"
         /clone_11d="tomato ovary, TAMU"
         /tissue_type="carpel"
         /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
         /lab_host="X11-Blue MRP"

```

```

/Note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLED - Tomato Carpel EST Library_ OligodT-primed and
directionally cloned cDNA in vector Lambda Zap II with 5'
and 3' ends located at the EcoRI and XhoI sites,
respectively."
BASE COUNT      201 a      120 c      155 g      191 t
ORIGIN
alignment_scores:
    Quality: 505.50      Length: 222
    Ratio: 3.159      Gaps: 1
    Percent Similarity: 72.072      Percent Identity: 46.396
alignment_block:
US-09-581-241-6 x A1486799 ..
Align seg 1/1 to: A1486799 from: 1 to: 667
280 GlnAspTyrLysSerSerValIleLeuValProThrLeuPheAla1 296
    ||::| ::::::::::::::
1 CAATAATTACAGAGTGCATAGGAGCCATTGTGCCACTTATTTGGC 50
296 eleuAsnArgSerGluLeuLeuAspLysTyrAspLeuSerAsnLeuVal 313
    :: ::::::::::::::
51 TATTGCTAAGAGCTCTATGCTGATTAATGATTAATGATTCATCGTAAGA 100
313 IuLeAlaSerGlyAlaProLeuSerLysGluIleGlyAlaVal 329
    :: ::::::::::::::
101 CCGTATGTCTGGGCTGCACCTTTAGAAAACCTTGAAACACTGTT 150
330 AlaArgArgPheAsnLeuProGlyValArgGlnGlyTyrGlyLeuThr 346
    :: ::::::::::::::
151 CGAGCCAAATTTCTTAATGCTTAACCTTGTCACAGTTACGATGACAGA 200
346 uThrThrSerAlaIleIle.....ThrProGluGlyA 358
    |::| ::::::::::::::
201 AGCCGACCAAGTGTGCTATGCTTGGCATTTGCTAAAGAACCTTGG 250
358 spAspLysProGlyAlaSerGlyLysValValProLeuPheLysAlaLys 374
    :: ::::::::::::::
251 AATTTAAATCTGGAGCATGTGGACTGTTGTAAGAAATGCTCGATGAGAA 300
375 ValIleAspLeuAspThrLysLysThrLeuGlyProAsnArgArgGly 391
    ::::::::::::::
301 ATGTGATCTCTGATCTGATGCAATCTCTCCCTAGGACCAATCTGGAGA 350
391 uValGlyValLysGlyProMetLeuMetLysGlyTyrValAspAsnPro 408
    |::| ::::::::::::::
351 GATTTGTATAGAGGCGCATCAATTATGAAAGTTACCTAAATGATCCAG 400
408 IuAlaThrArgGluIleIleAspGluGlyTyrLeuHisThrArgAsp 424
    ||||| ::::::::::::::
401 AGGCCCTACCGGACAAATAGACAAAGAAAGTGTTAATATACGAGCGCAT 450
425 IleGlyTyrTrpAspGluGluHisPhePheIleValAspArgLeu 441
    ||||| ::::::::::::::
451 ATGGTTATATTGACATCGATGATGAGCTTTTCATTGCGATCGTTTAA 500
441 sSerLeuIleLysTyrLysGlyTyrGlnValProProAlaGluLeuGlu 458
    |::| ::::::::::::::
501 GGAATTTGATTAATAATACAAAGATTCAGTGGCTCTCTGTAAGTTGAG 550
458 erValLeuGluGlnHisProAsnIlePheAspAlaGlyValAlaGlyVal 474
    ::::::::::::::
551 CCTTCTCTCATCTATCCCAATATTTTCAGATGCTGCTGTGTTCAAG 600
475 ProAspProIleAlaGlyLysLeuProGlyAlaValAlaValLeuLys 491
    ||| ::::::::::::::
601 AAAGACGACGACGAGAGAGATCTCTGCTGCTTTGTTGTCAGATCAAA 650
491 sGlyLysSerMetThr 496
    |::| ::::::::::::::

```

```
seq_name: gb_est2:B6414540  
seq_documentation_block:  
LOCUS B6414540      876 bp    mRNA                      EST            13-MAR-2001  
DEFINITION HYSMEX0002N03f Hordeum vulgare testa/pericarp EST library  
HVEDNA0013 (normal) Hordeum vulgare cDNA clone HYSMEK0002N03f, mRNA  
sequence.  
ACCESSION B6414540  
VERSION B6414540  
KEYWORDS B6414540.1 GI:13320091  
SOURCE barley.  
ORGANISM Hordeum vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
; Triticeae; Hordeum.  
REFERENCE 1 (bases 1 to 876)  
AUTHORS Wang,R., Close,T.J., Kleinhofs,A., Wise,R., Begun,D., Fritsch,D., Yu  
,Y., Anderson,H., Dale,J., Henry,D., Kednodle,S., Palmer,M., Rambo  
,T., Sasaki,C., Schwartzbeck,J., Simons,J., Chol,D.W., Main,D. and  
Wood,T.  
TITLE Development of a genetically and physically anchored EST resource  
for barley genomics  
JOURNAL Unpublished (2000)  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: AATTAACCCTCACTAAAGG  
High quality sequence stop: 847.  
Location/Qualifiers  
    1..876  
        /organism="Hordeum vulgare"  
          /cultivar="Morex"  
          /db_xref="taxon:4513"  
          /clone="HVSMEK0002N03f"  
          /clone_lib="Hordeum vulgare testa/pericarp EST library  
HVCDNA0013 (normal)"  
          /tissue_type="testa/pericarp"  
          /lab_host="TJC121"  
          /note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI; For  
more details on library preparation and sequence analysis  
see http://www.genome.clemson.edu/projects/barley/. To  
order a clone see http://www.genome.clemson.edu/orders"  
BASE COUNT      248 a       187 c       209 g       232 t  
ORIGIN
```

```
alignment_scores:
Quality:       505.50          length:        215
Ratio:         3.082           Gaps:           4
Percent Similarity: 76.279     Percent Identity: 49.767

alignment_block:
US-09-581-241-6 x B6414540 ..

Align seg 1/1   to: B6414540 from: 1 to: 876

323 TysGLUllEGlyGUAlAValAlArGARphAsnLnuPProGLyVALar 339
+++++::: |||||:::|||||::: |||||::: |:::
6 AGGGAGCTCATGTGAGTGCGCCAGAAGACTTCGCCGAGCGAGATCG 55
339 ggGlnGYTYrGlyLeuThrgLUthrTSerAlalAlellelFrhPrOG 356
|||||||:::|||||||::: |||||::: |:::
56 TCAGGGCTATTGATGACTGAGACTGTGGG.....ATCATATFCCTGG 99

356 luGlYspasplyspro..... GlyAlaserGlyLyssval 367
|||:::|||||::: |||||::: |||||::: |:::
100 AAATAACCAAGGAAGGACAAGCTGCATTTGGGTCATACCGGACGACTT 149
```

[illegible]

117 CCGAGCAGGTGGTGGCTATGCTTGGCATTTGCTAAAGAACCTTCGAA 166
 359 AsplysProglialaserglyLysValValProleuphetysalalysVa 375
 167 ATTAATCTGGAGCATGTGGACTGTTGTAGAAATGCTCGATGAATAAT 216
 375 lileaspleusprthlyslsThrluenglyProasnargraglygluv 392
 217 TGTGATCTGCTGATGACTGCAATATCTCCCTAGAGAACCAATCTGAGAGA 266
 392 alycysValysglYprometleumetlyslYTYrValalaspasProglu 408
 267 TTTGTATAGAGCGCATCAAAATTATGAAGGTTACCTAAAGATCCACAG 316
 409 AlathrargluilleleaspluenglYThrluenglyHstnrglyAspII 425
 317 GCCACTAGCGGACAAATAGACAAAGAGGTTGTTACATACGGCGGATAT 366
 425 eglYTYrTYrAspgluLysHlsPhepheleValasparleulys 442
 367 TGTATATATTGATCATGATGATGAGCTTTTCATGTGATGCTTAAAG 416
 442 erleuilelyslYrlyslYTYrGlnValProproalaglulengluser 458
 417 AATTGATTAATAACAAAGATTTCAAGTGGCTCTGCTGAACTTCAACCC 466
 459 ValleuenglnHlsProasnIlePheaspalaglYValalaglYValPr 475
 467 CTCTTCCTCAATCATCCCAATTTTCAGATGCTGCTGTTGTTCAATGAA 516
 475 OASProilealaglYgluLeuProglYAlaValAlaValleuLys 492
 517 ACACGAGCAAGCAGAGGAGTCTGCTGCTTTGTGTGACATCAATG 566
 492 lYlYsSerMetThrglyLysgluValMetAspTYrValAlaserglnVal 508
 567 GCTCCCAATCTACTGAGATGAAGTCAAGATTCATCAACANAGAGGTTG 616
 seq_name: gb_est1:BE454671
 seq_documentation_block: 686 bp mRNA EST 02-MAR-2001
 LOCUS BE454671
 DEFINITION HVSMEN0094P16f Hordeum vulgare 5-45 DAP spike EST library
 HVCN0094 (5 to 45 DAP) Hordeum vulgare cDNA clone HVSMEN0094P16f,
 mRNA sequence.
 BE454671
 VERSION BE454671.2 GI:13189489
 KEYWORDS EST.
 SOURCE
 ORGANISM
 Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Hordeum.
 1 (bases 1 to 686)
 Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
 Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo
 T., Sasaki,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and
 Wood,T.
 TITLE
 Development of a genetically and physically anchored EST resource
 for barley genomics
 JOURNAL
 Unpublished (2000)
 COMMENT
 On Jul 26, 2000 this sequence version replaced gi:9463485.
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seg primer: AATTAACTCTCACTAAAGG
 High quality sequence stop: 653.
 FEATURES
 Location/Qualifiers
 1..686

/organism="Hordeum vulgare"
 /cultivar="Morex"
 /db_xref="taxon:4513"
 /clone="HVSMEN0094P16f"
 /clone_lib="Hordeum vulgare 5-45 DAP spike EST library
 HVCN0094 (5 to 45 DAP)"
 /tissue_type="5-45 DAP Spike"
 /lab_host="SOLR"
 /note="Vector: lambdaZAP. Site 1: EcoRI; Site 2: XhoI. For
 more details on library preparation and sequence analysis
 see <http://www.genome.clemson.edu/projects/barley/> To
 order a clone see <http://www.genome.clemson.edu/orders>"
 BASE COUNT 156 a 212 c 203 g 115 t
 ORIGIN
 alignment_scores:
 Quality: 482.00 Length: 223
 Ratio: 2.957 Gaps: 2
 Percent Similarity: 73.094 Percent Identity: 43.049
 alignment_block:
 US-09-581-241-6 x BE454671 ..
 Align seg 1/1 to: BE454671 from: 1 to: 686
 314 lileaserglyLysValProleuserysluenglyLysValAla 330
 17 GTCAATGTCGGCGCGCGCCCATGGCAAGAGCTGACAGACCGCTCAT 66
 330 AARGAPheasleuProglYAlaArglnglyTYrGlyLeuThrgluT 347
 67 GGCAAGATCCCAACCGCGTCTGCGGAGGGGTACGGATGACGAGG 116
 347 hrThrsrAlaIlelelThrProgluLysAspLysPro..... 361
 117 CCGGCGCGGTGCTGCCATGCTGCTGCTGCCCAAGAGCGCTTCAG 166
 362GlyAlaserglyLysValValProleuphetysalalysVa 375
 167 GTCAAGTCCGGGTGCTGCGGACAGGTGTGCGCAAGCCGGGCTCAAGT 216
 375 lileaspleusprthlyslsThrluenglyProasnargraglygluv 392
 217 CGTCGACCCCGCACACCGCGGCTCTGCGCGGACAGCGCGCGCAGA 266
 392 alycysValysglYprometleumetlyslYTYrValalaspasProglu 408
 267 TGTGATCTGCGGAGCGAGATCAAGAGGTTACCTTAACGACCCAGAA 316
 409 AlathrargluilleleaspluenglYThrluenglyHstnrglyAspII 425
 317 TCACCAAGAACACCATGACACAGAGTGGCTGCGACACCGGACAT 366
 425 eglYTYrTYrAspgluLysHlsPhepheleValalaspargleulys 442
 367 CGGATCTCTCAGCAGCAGCAGAGATCTTCATGTTGACAGGCTCAAG 416
 442 erleuilelyslYrlyslYTYrGlnValProproalaglulengluser 458
 417 AGATCATCAAGTACAGGCTTCCAGGTGGCGCGGACGAGCTGAGAGCC 466
 459 ValleuenglnHlsProasnIlePheaspalaglYValalaglYValPr 475
 467 CTCTTCATCACCACCGAGATCAAGAGCCCGCGCTGATGCTGAA 516
 475 OASProilealaglYgluLeuProglYAlaValAlaValleuLys 492
 517 GCACATCTTCTGCGAAGTCCGATGCTCTTATGCGGATCGAAG 566
 492 lYlYsSerMetThrglyLysgluValMetAspTYrValAlaserglnVal 508
 567 GCTTGAATCAACGAGATGACATCAAAATTCGTGCGCAAGAGGTT 616

```

509 SerAsnAlaThrLeuArgGlyGlyValArgPheValaspGluValPro 525
      |||||..... ||| |||:::|||:::|||
617 GTTTCACAGAGATCCAC...MAGGTCTTTTCACGCATCAATCCC 663
      |||::: :||| |
825 OlysgLyLeuThrGlyLys 531
      |||::: :||| |
864 CAAGAACCCTTGACGCCAA 682

seq_name: gb_est2.BF623201

seq_documentation_block:
LOCUS       BF623201          817 bp            mRNA                    EST           21-FEB-2001
DEFINITION  HVSMEa001SK22f Hordeum vulgare seedling shoot EST library
            HVCNMA0001 (Cold stress) Hordeum vulgare cDNA clone HVSMEa0015K22f,
            mRNA sequence.
ACCESSION   BF623201
VERSION     BF623201.2  GI:13085328
KEYWORDS    EST.
SOURCE      barley.
ORGANISM    Hordeum vulgare
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
REFERENCE   1 (bases 1 to 817)
AUTHORS     Wang,R., Close,T.J., Kleinholz,A., Wise,R., Begum,D., Fritsch,D., Yu
            ,Y., Anderson,H., Dale,J., Henry,D., Kennode,S., Palmer,M., Rambo
            ,T., Sask,C., Schwartzbeck,J., Simons,J., Choi,D.W., Main,D. and
            Wood,P.
TITLE        Development of a genetically and physically anchored EST resource
             for barley genomics
JOURNAL     Unpublished (2000)
COMMENT     On Dec 18, 2000 this sequence version replaced gi:11887031.
Contact: Wing RA
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel.: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: AATTAAACCCCTGACTAAGCG
High quality sequence stop: 772.

FEATURES
         Source               1..817
                                /organism="Hordeum vulgare"
                                /cultivar="Morex"
                                /db_xref="taxon:4513"
                                /clone="HVSMEa0015K22f"
                                /clone_lib="Hordeum vulgare seedling shoot EST library
                                HVCNMA0001 (Cold stress)"
                                /tissue_type="Seedling shoot"
                                /lab_host="TJC121"
                                /note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI. For
                                more details on library preparation and sequence analysis
                                see http://www.genome.clemson.edu/projects/barley/To
                                order a clone see http://www.genome.clemson.edu/orders"

BASE COUNT      180 a          240 c          242 g          155 t

ORIGIN
alignment_scores:
Quality: 481.50              Length: 272
Ratio: 2.508                  Gaps: 3
Percent Similarity: 70.588    Percent Identity: 36.765

alignment_block:
US-09-581-241-6 x BF623201 ..

Align seg 1/1 to: BF623201 from: 1 to: 817

266 MetLeuThrLysPheAspGlnGluThrPheLeuLysThrLeuGlnAspTy 282
      ::::: ||| |||::: ::::::::::::::: ::::: ::
2 ATCATGCGCAAGGTGACGACGCGCGCGTGGTGGCGCTTGTCGCGCGCA 51

```

seq_documentation_block:

LOCUS BF064865 806 bp mRNA EST 09-MAR-2001
DEFINITION HV.CEBD0021.D12f Hordeum vulgare seedling green leaf EST library
HVCNDNA0005 (Erysiphe infected & control) Hordeum vulgare cDNA clone
HV_CEBD0021.D12f, mRNA sequence.

ACCESSION BF064865

VERSION BF064865.2 GI:13266207
 KEYWORDS EST.
 SOURCE Hordeum vulgare
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 1 (bases 1 to 806)
 REFERENCE
 1 (bases 1 to 806)
 Wang, R., Close, T.J., Kleinbols, A., Wise, R., Begum, D., Frisch, D., Yu
 Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo
 T., Sasaki, C., Schwartzbeck, J., Simmons, J., Choi, D.W., Main, D. and
 Wood, T.
 Development of a genetically and physically anchored EST resource
 for barley genomics
 Unpublished (2000)
 JOURNAL
 On Oct 17, 2000 this sequence version replaced gi:10841504.
 COMMENT
 Contact: Ming RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: AATTAACTCCTCACTAAAGG
 High quality sequence stop: 759.
 Location/Qualifiers
 1. 806
 /organism="Hordeum vulgare"
 /cultivar="C116151 (M186)"
 /db_xref="taxon:4513"
 /clone_lib="HV_CBD0021D12f"
 /clone="HV_CBD0021D12f"
 library HVCDNA0005 (Erysiphe infected & control)"
 /tissue_type="seedling green leaf"
 /lab_host="SOLR"
 /note="Vector: lambdaZAP, Site_1: EcoRI, Site_2: XhoI. For
 more details on library preparation and sequence analysis
 see http://www.genome.clemson.edu/projects/barley/ To
 order a clone see http://www.genome.clemson.edu/orders"
 BASE COUNT 185 a 201 c 239 g 181 t
 ORIGIN
 alignment_scores:
 Quality: 479.00 Length: 211
 Ratio: 3.051 Gaps: 3
 Percent Similarity: 74.408 Percent Identity: 45.972
 alignment_block:
 US-09-581-241-6 x BF064865 ..
 Align seg 1/1 to: BF064865 from: 1 to: 806
 340 GlnGlyTyrGlyLeuThrGluThrThrsAlaIleIleIleThrProG1 356
 |||||||
 3 CAGGGGTACGGGCTGACCGAGACGACGCGCTCCAGCTCAGCTC 52
 356 uG1AspAsp.....LysProGlyAlaSerGlyLysValValProLeuP 371
 :|||
 53 GGCCGAGAGAGCCGCCGCTACGGGACGCGGCTCTCTGCGCCCAACA 102
 371 helysAlaIleValIleAspLeuAspThrLysLysThrLeuGlyProAsn 387
 :|||
 103 CCGAGGCGCAAGATGCTGACCGGAGACGAGGCGCTGCGCGTGAAC 152
 388 ArgArgGlyGlyValLysValLysGlyProMetLeuMetLysGlyTyrVa 404
 |||
 153 CGCACCGCGCGAGCTCTGATCAGGCGGACCTACGATGAAAGATACTT 202
 404 LAspAsnProGluAlaThrArgGluIleLeuAspGluGlyTyrPleuH 421
 |||
 203 CAAGACACAGAGGACACAGTCAAGTCAAGCGTCAAGCCCAAGATGCTCA 252

421 1stThrGlyAspIleGlyTyrTyrAspGluGluLysHisPheIleVal 437
 |||||
 253 AGACCGGTGACCTATGCTACATAGACAGAGATGGATCTCTTCGGGG 302
 438 AspArgLeuLysSerLeuIleLysTyrLysGlyTyrGluValProProAl 454
 |||||
 303 GACCGTCTGAAGAGTTGATCAAAATCAAAAGCTATCAGTGCCTCCAC 352
 454 acIleuLeuGluSerValLeuLeuGluHisProAsnIlePheAspAlaGly 471
 |||||
 353 AGAGTTGGAGCTCTTCTGTACACCATCAAGGTTCCGATGCTGCG 402
 471 AlAlaGlyValProAspProIleAlaGlyLeuLeuProGlyAlaValAl 487
 |||||
 403 TTATTCCTTCCACAGACAGAGAGTGTGTCAGTTCCTCCATGGCTATGTC 452
 488 ValLeuLysLysGlyLysSerMetThrGlyLysGluValMetAspTyrVa 504
 |||||
 453 GTGAGAGAAAGAGGAGCAATTTGTCAGCCAGAGAGTGTAGATGCTGT 502
 504 lAlaSerGluValSerAsnAlaLysArgLeuArgGlyValAlaArgPheV 521
 |||||
 503 GCGGAACACAGTACGACCTTACACAGAGAGGTGAG...AAGTGGCGTTG 549
 521 AlAspGluValProLysGlyLeuThrGlyLysIleAspGlyAlaIle 537
 |||||
 550 TCACAGACATCCCAAGATGCGTCTCGCAAGATATGTAGAGAG..... 593
 538 ArgGluIleLeuLysLysProValAlaLysMet 548
 |||||
 594 ...GATCTTATCAAGCTCGCAGCTCCCAAGCTG 623
 seq_name: gb_est2:BB820931
 seq_documentation_block:
 LOCUS BB820931 772 bp mRNA EST 24-MAY-2001
 DEFINITION GM700013A20B2 Gm-r1070 Glycine max cDNA clone Gm-r1070-5044 3',
 mRNA sequence.
 ACCESSION BB820931
 VERSION BB820931.1 GI:10253165
 KEYWORDS EST.
 ORGANISM soybean.
 SOURCE Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 772)
 REFERENCE
 1 (bases 1 to 772)
 Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Corryell, V.,
 Expelling, J., Raph, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.
 A Functional Genomics Program for Soybean (NSF 9872565)
 Unpublished (1999)
 Other-ESTs: AW458737 corresponding to Gm-cl016-6629 (5')
 Contact: Vodkin, L.O., PI, A Functional Genomics Program for
 Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional
 Genomics
 University of Illinois
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582
 Email: l.vodkin@uiuc.edu
 This clone is available through: Genome Systems, Inc. 4633 World
 Parkway Circle St. Louis, Missouri 63134. For further information
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or info@genome
 systems.com web site: www.genomesystems.com
 Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.
 Location/Qualifiers
 1. 772
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="Gm-r1070-5044"

/clone.lib="Gm-r1070"
 /note="The library Gm-r1070 is a sequence-driven, rerecked set of 9,216 clones selected from cDNA libraries from various tissues and stages of development of soybean that represent 2,639 sequences from immature cotyledons, 1,770 from immature seed coats, 3,938 from flowers, and 869 from young pods. The 5' ESTs of the source clones from the different libraries was used to select singletons, or a representative of each contig, which were rerecked to form library Gm-r1070. The cDNA clones of the rerecked Gm-r1070 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the Laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://www.cbic.umn.edu/research/projects/soybean/index.html. Rerecking was performed by Genome Systems, St. Louis, http://www.genomesystems.com, and 3 sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, http://www.life.uiuc.edu/biotech/keck.html. Note: The corresponding 5' EST from each clone in the Gm-r1070 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Genome Systems clone ID of the original cDNA library that is also listed under 'OTHER EST'."

BASE COUNT 195 a 190 c 143 g 214 t 30 others
 ORIGIN

alignment_scores:
 Quality: 477.50 Length: 243
 Ratio: 2.776 Gaps: 5
 Percent Similarity: 70.782 Percent Identity: 43.210

alignment_block:
 US-09-581-241-6 x BE820931/rev ..

Align seg 1/1 to reverse of: BE820931 from: 1 to: 772

```

307 AspleuSerAsnLeuValGluIleAlaSerGlyValAlaProLeuSerly 323
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
747 GATNNNNCGTCGCGANNNGATCGTTGCNNNGATCGCGGAGAGAN 698
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
323 sgluilegylgluAlaValAlaArgArpHeAsnLeuProGlyValArg 340
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
697 ANNNNNGATGAGGAGCATCAAGCTAGTCCCAAGCTCGTCATGTC 648
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
340 lnglytyrglyLeuThrGluThrThrSerAlaIleIle..IleThrPro 355
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
647 AGGCAATACGGTTTAACGTAG.....TCTGCAGTTACCCGAAACCA 604
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
356 GluGlyAspAspLysProGlyAlaSerGlyLysValAlaProLeuPhe 372
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
603 GAGAGGCAAAATCAGTGGGAGCAACAGTAACTGATACCAATAATAGA 554
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
372 sAlaLysValIleAspLeuAspThrLysLysThrLeuGlyProAsnArg 389
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
553 AGCCAAATATGTGAAACCAAGGAAACCATGTTCCGGTGAC 504
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
389 rglygluValCysValLysGlyProMetLeuMetLysGlyTyrValAsp 405
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
503 AAGGAGAACTCTGGATCAGAGACCTTATGTAAGAAGTTATTCTTGCT 454
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
406 AspProGluAlaThrArgGluIleIleAspGluLysTyrPheLysIsth 422
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
453 GACCAAAAGCAACTTCAGCAACTTTC...GTGATGGGTGTTAAAGAC 407
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
422 rglAspIleGlyTyrTyrAspGluLysHisPhePheIleValAsp 439
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
406 TGGGACCTCTGTATTTGATACAAAGTTCTTGTATGTTGATAGATA 357
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
439 rglLeuLysSerLeuIleLysTyrLysGlyValProProAlaGlu 455
  
```

```

|||||.....|||||.....|||||.....|||||.....|||||.....
356 GGTGAAGAAGTTGATTAATATACAAAGGCTACCAAGTGTGCTCCGAGAG 307
456 LeuGluSerValIleLeuGluHisProAsnIlePheAspAlaGlyValAl 472
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
306 CTAGAAAGTTACTCTATCATTCATTCAGAGTAATATGCTGACGTTAT 257
472 aglValProAspProIleAlaGlyLysLysLeuProGlyAlaValAla 489
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
256 TCCATTACCTGATGATGATGACGAGCCAAAGTCCCATGCGCTTTGTGTA 207
489 euLysLysGlyLysSerMetThrGluLysGluValMetAspTyrValAla 505
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 GACAACCTCAAAAGTTCCCTTGGCCGACAGAAATTATGATTGTGGCC 157
506 SerIleValSerAsnAlaLysArgLeuArgGlyValAlaArgPheValas 522
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 AAGAGGTTTCACCGTACAAAGAAATAAGG...CGTATGACATTGTGTTAA 110
522 pglValProLysGlyLeuThrGlyLysIleAspGlyLysAlaIleArg 539
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
109 TTCCATTACCAAGAAATGCTGGGGAAGATP.....CTGAGGA 72
539 luIleLysLysProValAlaLysMet 548
    ||| ..... ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
71 AGGACTTAATAAACTGCGCTCTCTAGATTG 43
  
```

seq_name: gb_est1:AV58307

seq_documentation_block:

LOCUS AV58307 604 bp mRNA 07-SEP-2000

DEFINITION AV58307 Arabidopsis thaliana green siliques Columbia Arabidopsis

thaliana cDNA clone SQ094e05F 3', mRNA sequence.

ACCESSION AV58307

VERSION AV58307.1 GI:8729733

KEYWORDS EST.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 604)

AUTHORS Asanizu,E., Nakamura,Y., Sato,S. and Tabata,S.

TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation

of 12,028 non-redundant expressed sequence tags from normalized and

JOURNAL size-selected cDNA libraries

MEDLINE DNA Res. 7, 175-180 (2000)

COMMENT Contact: Erika Asanizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asanizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/Plant/.

FEATURES

1.604 Location/Qualifiers

source

1.604 /organism="Arabidopsis thaliana"

1.604 /strain="Columbia"

1.604 /db_xref="taxon:3702"

1.604 /clone="SQ094e05F"

1.604 /clone.lib="Arabidopsis thaliana green siliques Columbia"

1.604 /tissue_type="green siliques"

1.604 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 165 a 129 c 148 g 162 t

ORIGIN

alignment_scores:

Quality: 474.50 Length: 203

Ratio: 3.142 Gaps: 2

Percent Similarity: 74.384 Percent Identity: 47.291

alignment_block:

US-09-581-241-6 x AV558307

Align seg 1/1 to: AV558307 from: 1 to: 604

```
304 AsplysTyraAspleSerAsnLeuValGluIleAlaSerGlyGlyAlaIpr 320
      ::::::::::::::::::::: :::::::::::::::::::::
3   AATCTTATGATCTCTCCCTCGTATGATGCTTCTTCCGTCGACGCTCC 52
320 OlauSerLysGluIleGlyGluAlaValAlaArgArgPheAsnLeuProG 337
      ::::::::::::::::::::: :::::::::::::::::::::
53   TTTAGGTAAGAACTTCAAGATAGCTTTCGCGCGTCTCCCTCAAGCCA 102
337 LyValArgGlnGlyTyrglyLeuThrGluThrThrSerAlaIleIle 353
      ::::::::::::::::::::: :::::::::::::::::::::
103   TCCTTGGCCAGGATATGATATGACGAGGAGGACGACAGTGTGTCATG 152
354 Thr.....ProGluGlyAspAspLysProG1 362
      :::::::::::::::::::::
153   AGCCTTGCGTTGCTTAAGAACCGATACG.....ACAAATCAGG 193
362 YAlaSerGlyLyValValProLeuPheLysAlaLysValIleAspLeuA 379
      :::::::::::::::::::::
194   TTCAATGTGGGAGCTGTGTCGCAAGACGAGCTTAAGAGTGTACCTTG 243
379 sPThrLysLysThrLeuGlyProAsnArgArgGlyGluValCysValLys 395
      :::::::::::::::::::::
244   AGACACGCTCTCTCTGTGATACAAACCAACCTGAGATTGTATCCGC 293
396 GlyProMetLeuMetLysGlyTyValAspAsnProGluAlaThrArg1 412
      :::::::::::::::::::::
294   GGTCAACAGATCATGAAGAGTACTGACGATCCAGAACCCACTTCAGC 343
412 uIleIleAspGluGluGlyTyTrpLeuHisThrGlyAspIleGlyTyTyra 429
      :::::::::::::::::::::
344   AACATATAGACGAAGAAGTGTGCTTCACACAGAGACATGATATGTG 393
429 sPgIuGluLysHisPhePheIleValAspArgLeuLysSerLeuIleLys 445
      :::::::::::::::::::::
394   ATGAAGATGATGATTTTCATGTGTGATGACTTAAGAAGTCATCAG 443
446 TyTrLysGlyTyrglyValProProAlaGluLeuGluSerValLeuLeuG1 462
      :::::::::::::::::::::
444   TTCAAAGGCTTTCAGTCCCTCCAGCTGAGCTGAGAGTTTGCTAATCAA 493
462 nHisProAsnIlePheAspAlaGlyValAlaGlyValProAspProIleA 479
      :::::::::::::::::::::
494   TCACCATTCATATGCCGATGACGCTGTGTTCCGCAAAATGATGAGTCG 543
479 IacGlyGluLeuProGlyAlaValValValLeuLysGlyLysSerMet 495
      :::::::::::::::::::::
544   CTGGGGAAGTCCGCTGCTTTCGTGTGCGATCAATGGAATGATATC 593
496 ThrGluLys 498
      :::::::::::::::::::::
594   ACTGAAGAA 602
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 01:53:50 ; Search time 25.81 Seconds

(without alignments)
778.471 Million cell updates/sec

Title: US-09-581-241-4

Perfect score: 2823
Sequence: 1 MEMMENDEIVYGPPEPPI.....TKRIDKAIKREILKKPVAKM 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2814	99.7	548	1	LUCI_LUCIA
2	2684	95.1	548	1	LUCI_LUCR
3	1945.5	68.9	550	1	LUCI_PHOY
4	768.5	27.2	544	1	ACLI_PENCR
5	760.5	26.9	544	1	ACLI_PENCR
6	758.5	26.9	547	1	ACLI_TOBAC
7	748.5	26.5	542	1	ACLI_TOBAC
8	747	26.5	553	1	ACLI_TOBAC
9	734.5	26.0	543	1	ACLI_VANPL
10	733	26.0	543	1	ACLI_SOLTU
11	732	25.9	561	1	ACLI_ARYTH
12	727.5	25.8	545	1	ACLI_ORYSA
13	726	25.7	562	1	ACLI_SOLTU
14	720.5	25.5	561	1	ACLI_ARYTH
15	712	25.4	536	1	ACLI_ARYTH
16	689.5	24.2	569	1	ACLI_ORYSA
17	674	23.9	537	1	ACLI_ORYSA
18	587.5	20.8	293	1	ACLI_SOYBN
19	572.5	20.3	560	1	ACLI_BACSU
20	569	20.2	561	1	ACLI_BACSU
21	481	17.0	562	1	ACLI_ECOLI
22	463	16.4	543	1	ACLI_HAELN
23	461	16.3	572	1	ACLI_YEAST
24	444	15.7	548	1	ACLI_BACSU
25	427.5	15.1	522	1	ACLI_ECOLI
26	368	13.0	698	1	ACLI_HUMAN
27	366	13.0	683	1	ACLI_HUMAN
28	359	12.7	699	1	ACLI_HUMAN
29	357	12.6	699	1	ACLI_HUMAN
30	351.5	12.5	699	1	ACLI_HUMAN
31	350	12.4	699	1	ACLI_HUMAN
32	346.5	12.3	683	1	ACLI_HUMAN
33	338	12.0	698	1	ACLI_HUMAN

34	334.5	11.8	539	1	DBE_BACSU
35	325	11.5	711	1	ICPD_HUMAN
36	324.5	11.5	670	1	ICPD_RAT
37	323.5	11.5	6359	1	BACC_BACLI
38	321	11.4	697	1	ICPF_HUMAN
39	317.5	11.2	694	1	ICPF_YEAST
40	312	11.1	697	1	ICPF_RAT
41	311.5	11.0	543	1	VIBE_VIBCH
42	308.5	10.9	711	1	ICPD_MOUSE
43	308.5	10.9	744	1	ICPD_YEAST
44	306	10.8	5255	1	BACA_BACLI
45	305.5	10.8	700	1	ICP1_YEAST

ALIGNMENTS

RESULT	ID	STANDARD	PRT	548 AA.
1	LUCI_LUCIA			
1	LUCI_LUCIA			
AC	001158:			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).			
OS	Luciola lateralis (Firefly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;			
OC	Elateriformia; Cantharoidae; Lampyridae; Luciola.			
OX	NCBI_TaxID=7052;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-92305054; PubMed-1610896;			
RA	Tatsumi H., Kajiyama N., Nakano E.;			
RT	"Molecular cloning and expression in Escherichia coli of a cDNA clone			
RT	encoding luciferase of a firefly, Luciola lateralis.";			
RL	Biochim. Biophys. Acta 1131:161-165(1992).			
CC	- FUNCTION: PRODUCES GREEN LIGHT.			
CC	- CATALYTIC ACTIVITY: LUCIFERIN + O(2) + ATP = OXIDIZED LUCIFERIN +			
CC	CO(2) + H(2)O + AMP + PYROPHOSPHATE + LIGHT.			
CC	- CORFACTOR: REQUIRES MAGNESIUM.			
CC	- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).			
CC	- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME			
CC	FAMILY.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement and its license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR	EMBL: X66919; CAA47358.1; -
DR	PIR: S23437; S23437.
DR	HSSP: P08659; LUCI.
DR	Interpro: IPR00873; AMP-bind.
DR	Pfam: PF00501; AMP-binding; 1.
DR	PROSITE: PS00342; MICROBODIES_CTER; 1.
DR	PROSITE: PS00455; AMP-BINDING; 1.
KW	Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
FT	Peroxisome.
FT	SITE
SEQUENCE	548 AA: 60125 MW; ACQ2P9320BB6DA6 CRC64;

Query Match 99.7%; Score 2814; DB 1; Length 548;
Best Local Similarity 99.6%; Pred. No. 1.4e-188;
Matches 546; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	1	MEMMENDEIVYGPPEPPIEGSGAGOLRKTYMDRYAKGATAFNALTGVDYTAETYE 60
DB	1	MEMMENDEIVYGPPEPPIEGSGAGOLRKTYMDRYAKGATAFNALTGVDYTAETYE 60

```

QY 61 KSCCLGKALKNYGLVNDGRIALCSENCEEEFFIPVLAGLFGVGAAPNEITTLRELHSL 120
DB 61 KSCCLGKALKNYGLVNDGRIALCSENCEEEFFIPVLAGLFGVGAAPNEITTLRELHSL 120
QY 121 GISKPTIVFSSKRGKLDKVTIVOKTVAIKTIVILDSKVYRGYOSMDNFIKKNTPGQFKG 180
DB 121 GISKPTIVFSSKRGKLDKVTIVOKTVAIKTIVILDSKVYRGYOSMDNFIKKNTPGQFKG 180
QY 181 SSRKTIVANRKEQVALLMNSSGSTGLPKGVQLTHENLVTRSHARDPIYGNOVSPGTAIL 240
DB 181 SSRKTIVANRKEQVALLMNSSGSTGLPKGVQLTHENLVTRSHARDPIYGNOVSPGTAIL 240
QY 241 TVPFEHGFHGFMTLLGYLTCGFRIVMLTKPEDEFTFLKLDQYKSSVILVPTLPAIILNS 300
DB 241 TVPFEHGFHGFMTLLGYLTCGFRIVMLTKPEDEFTFLKLDQYKSSVILVPTLPAIILNS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETTSALITPEGDDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETTSALITPEGDDK 360
QY 361 PGASGKVVPLFKAVIDLDTRKKTIGPNRGEVCYKGPMLMGVYDNEPATREITIDEGMWL 420
DB 361 PGASGKVVPLFKAVIDLDTRKKTIGPNRGEVCYKGPMLMGVYDNEPATREITIDEGMWL 420
QY 421 HTGDIGYDEKHFIVDRKSLIKYGYOVPAPLESVLLQHNIPDAGVAGVPPDIAG 480
DB 421 HTGDIGYDEKHFIVDRKSLIKYGYOVPAPLESVLLQHNIPDAGVAGVPPDIAG 480
QY 481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNARKRLGVRVDEVPKGLTGKIDKAIKEI 540
DB 481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNARKRLGVRVDEVPKGLTGKIDKAIKEI 540
QY 541 LKRPVAKM 548
DB 541 LKRPVAKM 548

RESULT 2
LUC1_LUCR STANDARD; PRT; 548 AA.
AC P13129;
DB 01-JAN-1990 (Rel. 13, Created)
DB 01-JAN-1990 (Rel. 13, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).
OS Luciola cruciata (Japanese firefly) (Genji firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharidoidea; Lampyridae; Luciola.
OC NCBI_TaxID=7051;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89326143; PubMed=2473944;
RA Masuda T., Tatsumi H., Nakano E.;
RT "Cloning and sequence analysis of cDNA for luciferase of a Japanese
firefly, Luciola cruciata.";
RT Gene 77:265-270(1989).
CC -1- FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 544 NM.
CC -1- CATALYTIC ACTIVITY: LUCIFERIN + O(2) + ATP -> OXIDIZED LUCIFERIN +
CO(2) + H(2)O + AMP + PYROPHOSPHATE + LIGHT.
CC -1- COFACTOR: REQUIRES MAGNESIUM.
CC -1- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

```

CC EMBL: M26194; AAA29135.1;
DR PIR: JS0181; JS0181.
DR HSSP: P08659; 1IC1.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00342; MICROBODIES_CTER; 1.
DR PROSITE: PS00455; AMP-BINDING; 1.
KW Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
KM Peroxisome.
FT SITE 546 548 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 548 AA; 60017 MW; 2052D6189E79109F CRC64;

Query Match 95.1%; Score 2684; DB 1; Length 548;
Best Local Similarity 93.4%; Pred. No 1.6e-179;
Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

QY 1 MENNENDENTIVYGPPEYPIEESGAGQALRKMDRYAKLGAIAFTNALGVDTYAEYLE 60
DB 1 MENNENDENTIVYGPKPPIEESGAGTQLRKMYERAKLGAIAFTNAVTVGVSYAEYLE 60
QY 61 KSCCLGKALKNYGLVNDGRIALCSENCEEEFFIPVLAGLFGVGAAPNEITTLRELHSL 120
DB 61 KSCCLGKALKNYGLVNDGRIALCSENCEEEFFIPVLAGLFGVGAAPNEITTLRELHSL 120
QY 121 GISKPTIVFSSKRGKLDKVTIVOKTVAIKTIVILDSKVYRGYOSMDNFIKKNTPGQFKG 180
DB 121 GISKPTIVFSSKRGKLDKVTIVOKTVAIKTIVILDSKVYRGYOSMDNFIKKNTPGQFKG 180
QY 181 SSRKTIVANRKEQVALLMNSSGSTGLPKGVQLTHENLVTRSHARDPIYGNOVSPGTAIL 240
DB 181 SSRKTIVANRKEQVALLMNSSGSTGLPKGVQLTHENLVTRSHARDPIYGNOVSPGTAIL 240
QY 241 TVPFEHGFHGFMTLLGYLTCGFRIVMLTKPEDEFTFLKLDQYKSSVILVPTLPAIILNS 300
DB 241 TVPFEHGFHGFMTLLGYLTCGFRIVMLTKPEDEFTFLKLDQYKSSVILVPTLPAIILNS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETTSALITPEGDDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVROGYGLTETTSALITPEGDDK 360
QY 361 PGASGKVVPLFKAVIDLDTRKKTIGPNRGEVCYKGPMLMGVYDNEPATREITIDEGMWL 420
DB 361 PGASGKVVPLFKAVIDLDTRKKTIGPNRGEVCYKGPMLMGVYDNEPATREITIDEGMWL 420
QY 421 HTGDIGYDEKHFIVDRKSLIKYGYOVPAPLESVLLQHNIPDAGVAGVPPDIAG 480
DB 421 HTGDIGYDEKHFIVDRKSLIKYGYOVPAPLESVLLQHNIPDAGVAGVPPDIAG 480
QY 481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNARKRLGVRVDEVPKGLTGKIDKAIKEI 540
DB 481 ELPGAVVVLKKGKSMTEKEVMDYVASOVSNARKRLGVRVDEVPKGLTGKIDKAIKEI 540
QY 541 LKRPVAKM 548
DB 541 LKRPVAKM 548

RESULT 3
LUC1_PROPY STANDARD; PRT; 550 AA.
AC P08659;
DB 01-JAN-1988 (Rel. 06, Created)
DB 01-JAN-1988 (Rel. 06, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).
OS Photinus pyralis (North American firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharidoidea; Lampyridae; Photinus.
OC NCBI_TaxID=7054;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=87144243; PubMed=3821727;
RA de Wet J.R., Wood K.V., Deluca M., Helinski D.R., Subramani S.;
RT "Firefly luciferase gene: structure and expression in mammalian
RL cells.";
RL Mol. Cell. Biol. 7:725-737(1987).
[2]
RN SUBCELLULAR LOCATION.
RP MEDLINE=87204117; PubMed=3554235;
RA Keller G.-A., Gould S., de Luca M., Subramani S.;
RT "Firefly luciferase is targeted to peroxisomes in mammalian cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:3264-3268(1987).
[3]
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RP MEDLINE=96398615; PubMed=8805533;
RA Conti E., Franks N.P., Brick P.;
RT "Crystal structure of firefly luciferase throws light on a
RL superfamily of adenylate-forming enzymes.";
RL Structure 4:287-298(1996).
[4]
RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RP MEDLINE=99007339; PubMed=9788915;
RA Franks N.P., Jenkins A., Conti E., Lieb W.R., Brick P.;
RT "Structural basis for the inhibition of firefly luciferase by a
RL general anesthetic.";
RL Biophys. J. 75:2205-2211(1998).
CC -1- FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 562 NM.
CC -1- CATALYTIC ACTIVITY: LUCIFERIN + O(2) + ATP = OXIDIZED LUCIFERIN +
CC CO(2) + H(2)O + AMP + PYROPHOSPHATE + LIGHT.
CC -1- COFACTOR: REQUIRES MAGNESIUM.
CC -1- SUBCELLULAR LOCATION: PEROXISOMAL.
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15077; AAA29795.1; -
DR EMBL; X84848; CAA59283.1; -
DR EMBL; U03687; AAA03561.1; -
DR EMBL; U89934; AAB64396.1; -
DR EMBL; U89935; AAB64399.1; -
DR PIR; A26772; A26772.
DR PDB; 1LC1; 26-MAR-97.
DR PDB; 1BA3; 11-NOV-98.
DR InterPro; IPR000873; AMP-Bind.
DR Pfam; PF00501; AMP-binding.1.
DR PROSITE; PS00342; MICROBODIES_CTR; 1.
DR PROSITE; PS00455; AMP-BINDING.1.
DR Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
DR Peroxisome; 3D-structure.
DR SITE 548 550 MICROBODY TARGETING SIGNAL.
DR SITE 548 550 E380RCEPD56ACDCE CMC6;
DR SEQUENCE 550 AA; 60745 MW; E380RCEPD56ACDCE CMC6;
Query Match 68.9%; Score 1945.5; DB 1; Length 550;
Best Local Similarity 68.2%; Pred. No. 4,2e-128;
Matches 369; Conservative 75; Mismatches 96; Indels 1; Gaps 1;
QY 4 MENENITYYGEPPYPIEESAGQAKRYMDRYAKL-GAIFTNALITGVDTYATYLEKS 62
DB 1 MEDANNIKGAPPEYLEDGTAGEDQLHRAKRYALVPGTIAFTAHIEVNITYAEYFEMS 60
QY 63 CCLGALKNYGLVVDGIALCEEFPIVLAGLFTGVGVAPEINELYTLRELVHSLGI 122
DB 61 VRLAEAMKRYGLNTRHRYVCSNLSQFFMRYLGLFTGVAVAPANDIYNERELINSANI 120
QY 123 SKPIVSSKGLDKVITVQKTYAIIKTIIVLDSKVDYRGYOSMDNFITKKNTPQGFKSS 182

```

```

DB 121 SQPTVFWFSKGLQKILANVQKLPILQKIIIMDSKTDYQFQSMYTFVSHLPFGFNEYD 180
QY 103 FETVENREKQVALLIMNSGSGTGRKQVULHENVLRSHARDPIYGNVSPGAILTY 242
DB 181 FPPESFDRDKTIALIMNSSGSTGLPRGVALPHRTACVRSNARDPIFGNIIIDTALISV 240
QY 243 VPFHNGFMTTLGLTGCFRIVMLTKPDEFELTLDQKSSVILVPTLPAIINRSEL 302
DB 241 VPFHNGFMTTLGLTGCFRIVMLTKPDEFELTLDQKSSVILVPTLPAIINRSEL 300
QY 303 LDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVROGYLTETTSALITIPEDDGRG 362
DB 301 IDKYDLSNLVEIASGAPLSKEIGAVARFNLPGVROGYLTETTSALITIPEDDGRG 360
QY 363 ASGKVPPLFKAVITDLDRTKTGTPNRRCGVCKGPMKMGVYNDPEAREITDEGWLHT 422
DB 361 AVGKVPPEFAKAVYDLDGKTLGVNQRELGVGRPMKMGVYNNPEATVALIDKGMWHS 420
QY 423 GDIGYDEEKHFPIYDRLSLTKYGYOVPPAPLESVILQHPNIFDAGVAGVDPPIAGEL 482
DB 421 GDIAWDEDEHFFIYDRLSLTKYGYOVPPAPLESVILQHPNIFDAGVAGVDPPIAGEL 480
QY 483 PGAVVVLKKGKSMTEKEVADYVAVSOVSNARKLRGVREVEPKGLTGKIDKAIRELIL 542
DB 481 PAAVVVLEHGKTMTEKEIYDVAVSOVTTAKLRGVVFVDEVPKGLTGKIDKAIRELIL 540
QY 543 K 543
DB 541 K 541
RESULT 4
ACLI_PENCHR STANDARD: PRT; 544 AA.
ID ACCL_PENCHR
AC P14912;
DR 01-APR-1990 (Rel. 14, Created)
DR 10-APR-1990 (Rel. 14, Last sequence update)
DR 20-AUG-2001 (Rel. 40, Last annotation update)
DE 4-COMMARATE-COA LIGASE 1 (EC 6.2.1.12) (4CL 1) (4-COMMAROYL-COA
DE SYNTHASE 1).
GN 4CL OR 4CL-1.
OS Petroselinum crispum (Parsley) (Petroselinum hortense).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
OX NCBI_TaxID=4043;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89005119; PubMed=3169018;
RA Lozoya E., Hoffmann H., Douglas C., Schulz W., Scheel D.,
RA Hahlbrock K.;
RT "Primary structures and catalytic properties of isoenzymes encoded by
RT the two 4-commarate:COA ligase genes in parsley.";
RL Eur. J. Biochem. 176:661-667(1988).
[2]
RN SEQUENCE OF 1-8 FROM N.A.
RP Douglas C., Hoffmann H., Schulz W., Hahlbrock K.;
RT "Structure and elicitor or U.V.-light-stimulated expression of two
RT 4-commarate:COA ligase genes in parsley.";
RL EMBL J. 6:1189-1195(1987).
CC -1- CATALYTIC ACTIVITY: ATP + 4-COMMARATE + COA = AMP + DIPHOSPHATE +
CC 4-COMMAROYL-COA.
CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -1- INDUCTION: TRANSCRIPTIONALLY ACTIVATED BY BOTH FUNGAL ELICITOR AND
CC ULTRAVIOLET IRRADIATION.
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```


QY 442 SLIKYGVPPAELESVLLQHPNIPDAGVAGVDPPIAGELPGAVVYLKKKSKTEKEM 501
DB 439 EIRYKGVAPAELEMLLHPPTISDAVPMDEKAGVPAVFTVNGFTTEBEIK 498
QY 502 DYVASOVNNAKRLRGVFEVDEVGKLTGKIDGKAIR 538
DB 499 QVSKQVFFYKRI-FRVFVDAIPKSPSGKILRRDLR 534

RESULT 6
ID 4CL1_TOBAC STANDARD; PRT; 547 AA.
AC 024145:
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE 4-COMARATE-COA LIGASE 1 (EC 6.2.1.12) (4CL 1) (4-COMAROYL-COA SYNTHASE 1).
GN 4CL1.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_Taxid=4097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96416441; PubMed=8819324;
RA Lee D., Douglas C.J.;
RT "Two divergent members of a tobacco 4-coumarate:coenzyme A ligase (4CL) gene family. cDNA structure, gene inheritance and expression, and properties of recombinant proteins.";
RL Plant Physiol. 112:193-205(1996).
CC -1- CATALYTIC ACTIVITY: ATP + 4-COMARATE + COA = AMP + DIPHOSPHATE + 4-COMAROYL-COA.
CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROANOIC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U50845; AAB18637.1; -
DR HSSP: P08659; ILCI.
DR Mendel: 24477; Nicla:1179;24477.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP-BINDING; 1.
KW ligase; Phenylpropanoid metabolism; Multigene family.
KW SEQUENCE 547 AA; 59842 MW; 154DF6D684E3F51B CRC64;

Query Match 26.9%; Score 758.5; DB 1; Length 547;
Best Local Similarity 35.1%; Pred. No. 1.9e-45;
Matches 189; Conservative 89; Mismatches 205; Indels 55; Gaps 12;

QY 32 YMDRAKAGATAFTN-----ALTGVD---YTYAEYLRKSC-CIGELAKNYGLVVDG 78
DB 25 YIPKHLPHSHSCFENISFSSRPCLINCANDOIITYAE-VELTCKKAVGLNKIGIQOKD 83
QY 79 IALSCENCEFFIPLVLAGLFGVAPATNEIYTLRELVHSLGSKPTIVSSKGLDKV 138
DB 84 TIMILLPSPEVFAPMGASVIGALSTMANPLFPAEVVQAKAKSSAIIITQSCFVGKV 143
QY 139 ITVQTVTAIKTIVLDSKVDYRGYSMDNFKTKNTPQG-----FGSSFKYEVN-R 190
DB 144 -----KDYASENDVKVICIDSAPEGCLHFSLTLQSDHEIPEVXIQ 184

QY 191 KEQVALINMSGSTGPKGVOLTHENLYTRFSHARDPIYGN-QVSPGATLTVPRHHGF 249
DB 185 PDVVALLPVSSGTTGLPKRGVMLHKGLTVSAQOVDGENANLYMSEDVLMCVLFLH-- 242
QY 250 GMFTTGLTYLTCGR-----IVMLTKFEDEFTLTKTDYKSCSYLVPTLPAIINRSELDK 305
DB 243 -IYSLNSILLGIRVGAAILIMQKFDIAFLEIDQKTKYSIDPFVPIYLAIAKAPIYDS 301
QY 306 YDISNLVETASGAPLSKEIGEAVARFNLPGVROGYGLTETSAILI-----TEBGDK 360
DB 302 YDLSRVTVWSGAAPLGELEDAVRFPFNAKIGQGYGTEAGPVLAMKLAKRPFPIK 361
QY 361 PGASGVVPLPFKAVTIDLTCKTLGPNRGEVGVSGPMLAKGYVONPEATREIDEQWL 420
DB 362 SGACGVVNRNAEMKTYDPDGTGSLPRNQPGELCRGDOIIMKGLINPEATRTTIDKEGL 421
QY 421 HTGDIGYDEEKHFTVDRLSLIKKGYPAPAELESVLLQHPNIPDAGVAGVDPPIAG 480
DB 422 HTGDIGFIDEDELFIYDRKLKELIKYKGFQVAPAEIEALLNHPNISDAVPMKDEQAG 481
QY 481 ELPGAVVYLKKKSKMTEKEMDYVASOVNNAKRLRGVFEVDEVGKLTGKIDGKAIR 538
DB 482 EVPAEVARVNSGSAITDEKDFISKQVIFYKRVK-RVFEVETVPKSPSGKILRRDLR 538

RESULT 7
ID 4CL2_TOBAC STANDARD; PRT; 542 AA.
AC 024146:
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 4-COMARATE-COA LIGASE 2 (EC 6.2.1.12) (4CL 2) (4-COMAROYL-COA SYNTHASE 2).
GN 4CL2.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_Taxid=4097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96416441; PubMed=8819324;
RA Lee D., Douglas C.J.;
RT "Two divergent members of a tobacco 4-coumarate:coenzyme A ligase (4CL) gene family. cDNA structure, gene inheritance and expression, and properties of recombinant proteins.";
RL Plant Physiol. 112:193-205(1996).
CC -1- CATALYTIC ACTIVITY: ATP + 4-COMARATE + COA = AMP + DIPHOSPHATE + 4-COMAROYL-COA.
CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROANOIC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U50846; AAB18638.1; -
DR HSSP: P08659; IBA3.
DR Mendel: 24478; Nicla:1179;24478.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP-BINDING; 1.
KW ligase; Phenylpropanoid metabolism; Multigene family.
KW SEQUENCE 542 AA; 59479 MW; CB5579AEDDFC003 CRC64;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae: euasterids I; Solanales; Solanaceae; Solanum.
 OC NCBI_TaxID=4113;
 RN [1]
 RX SEQUENCE FROM N.A.
 RA MEDLINE=91217100; PubMed=2022667;
 RA Becker-Andre M., Schulze-Lefert P., Hahlbrock K.;
 RA "Structural comparison, modes of expression, and putative cis-acting
 RT elements of the two 4-coumarate: COA ligase genes in potato.";
 RL J. Biol. Chem. 266:8551-8559(1991).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
 CC 4-COUMAROYL-COA.
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M62755; AAA33842.1; -
 DR PIR: A39827; A39827.
 DR HSSP: P08659; ILCI.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-BINDING; 1.
 DR Ligase; Phenylpropanoid metabolism; Multi-gene family.
 SO SEQUENCE 545 AA; 59619 MW; DEL83683B774BA71 CRC64;

Query Match 26.08; Score 734.5; DB 1; Length 545;
 Best Local Similarity 34.28; Pred. No. 9e-44;
 Matches 185; Conservative 93; Mismatches 202; Indels 61; Gaps 12;

QY 32 YMDRYAKGAIATFN-----ATGCGVDD--TYVAEYLEKSCCLDEALKNVGLVYDGR 79
 DB 23 YTPHPLPHSYCFENLSEFNSRPLLIDANDRITYVAEVELTSKRVAVGLKLGIDQKDT 82
 QY 80 IALSCNCEEEFFIVLAGLFTGVAPAPTEIYLTRELVHSLGSKIPYVESSSKGLDKVI 139
 DB 83 IMILLPNEEVPFAFGISYLGAISTMANPPLTFPAVYKQAKASAKIVITQACFACKV- 141
 QY 140 TVQKTVTAIKTIVILDSKVDRYQSDMNFK---KNTPOG-----FKGSSFKTVEV 188
 DB 142 -----KDY-----ATENDLKVICVDSVPEGCHFSFELIOSDEHEIPDV 179
 QY 189 N-RKQVALLINSSGSGTGLPKGYQLTHENLVTRFSHADPITYGN-QVSPGAILTVYFPH 246
 DB 180 KIOPDVALVYSSGTGLPKGYMLTHKGLTVYAQVQDGENALIMHSDVLYMCVLPLE 239
 QY 247 HGFMTLLGYLTGCFR---IVMLTKFEDETFKTLQDYKCSSVILVPLFLALINSEL 302
 DB 240 H---IYSLNSVLICALRGAILLMOKFDIAQLELLPKHVTITGPPFPIVLAIAASPL 296
 QY 303 LDKYLSMLVEIASGAPLSKEIGEAVARRNLPQGVROGYGLTETTSATIT-----TPEG 357
 DB 297 VDNYSLSVRYVMSGAPLGELEDAVARAKPPNKLGGYGMTGAPLACLAFAKEPF 356
 QY 358 DDKPASKVYVPLFAKYIDTDTKKTGPNRRGEVCGPMLMGYVNPATREIIDEE 417
 DB 357 DIKSGACGTIVANMAEMKIVDPDTGCSLPDRNGPELCINIGDOIMKGYLMDPRATARTI 416
 QY 418 GMLHTGIDGYDEKHFIVRLKSLIKYGYOYPAPELESVLLDHPNIPDAGVAGYDDP 477
 DB 417 GMLHTGIDGIDDDDELFLVORLKEILIKYGYOYPAPELESVLLDHPNIPDAGVAGYDDP 476
 QY 478 TAGELPGAVVYLKKKSKTEKEVMDYVAVSOYNAKRLGVRFEVDEYVKGTLGKIDGKAI 537

DB 477 OAGEVPVAFVNSNGSTTEDEVKDFISKQVIFYRKIK-RVFEVETVPKSPSKILRDL 535
 QY 538 R 538
 DB 536 R 536

RESULT 10
 ID 4CL3.ARATH STANDARD; PRT; 561 AA.
 AC 09S777;
 DT 20-AUG-2001 (Rel. 40; Created)
 DT 20-AUG-2001 (Rel. 40; Last sequence update)
 DT 20-AUG-2001 (Rel. 40; Last annotation update)
 DE 4-COUMARATE--COA LIGASE 3 (EC 6.2.1.12) (4CL 3) (4-COUMAROYL-COA
 DE SYNTHASE 3)
 GN 4CL3 OR AT1G65060 OR F1G16.6.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=99348176; PubMed=10417722;
 RA Ehling J., Buettner D., Wang Q., Douglas C.J., Somsich I.E.,
 RA Kombrink E.;
 RT "Three 4-coumarate:coenzyme A ligases in Arabidopsis thaliana
 RT represent two evolutionarily divergent classes in angiosperms.";
 RL Plant J. 19:9-20(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.D., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Elgu P., Feldhlyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremetska I., Kutz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
 RA Uteback T., Van Aken S., Vayberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820(2000).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
 CC 4-COUMAROYL-COA.
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF106087; AAD47194.1; -
 DR EMBL: AF106088; AAD47195.1; -
 DR EMBL: AC009360; AAF06039.1; -
 DR HSSP: P08659; ILCI.

DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 4-COMARATE--COA LIGASE 2 (EC 6.2.1.12) (4CL 2) (4-COMAROYL-COA
 DE SYNTHASE 2).
 GN 4CL2 OR 4CL-2.
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91217100; PubMed=2022667;
 RA Becker-Andre M., Schulze-Lefert P., Hahlbrock K.;
 RT "Structural comparison, modes of expression, and putative cis-acting
 RT elements of the two 4-coumarate: CoA ligase genes in potato.";
 RL J. Biol. Chem. 266:8551-8559(1991).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COMARATE + COA = AMP + DIPHOSPHATE +
 CC 4-COMAROYL-COA.
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC PIR: B39827; B39827.
 DR HSSP; P08659; ILCI.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 1.
 DR PROSITE; PS00455; AMP_BINDING; 1.
 KW Ligase; Phenylpropanoid metabolism; Multigene family.
 SQ SEQUENCE 545 AA; 59625 MW; 5481F0B0AFA39E0 CRC64;

Query Match 25.8%; Score 727.5; DB 1; Length 545;
 Best Local Similarity 34.0%; Pred. No. 2.8e-43;
 Matches 184; Conservative 93; Mismatches 203; Indels 61; Gaps 12;

QY 32 YMDRYAKGAIATFTN-----ALTGVD---YTYAEYLEKSCCLGALKNYGLVVDGR 79
 DB 23 YPKHPLHSGCFENLSEFNRPCLIDGANDRIYAYAEVLETSRKVAVGLKGLIQKDT 82
 QY 80 IALGENCEBEFFIPVLAFLGCVGAPNNEYITLRELHSLGICRPTVFSKGLDVI 139
 DB 83 IMILPNCPEVFARIGASYGALSTMANPJTTPAEVVKOKAKSAAKIVITQACRAGV 141
 QY 140 TYQKVTAKITVILDSVDRGYQSDMNFK---KMPDG-----FGSSFKTEYV 188
 DB 142 -----KDY-----AIENDLKVTCVDSAPCECAFSELIQSDHEITPDV 179
 QY 189 N-RKDOVALINSSGSLGPKGVOLTHENLVTRFSHARDPIYGN-OVSPGTAITLVVPH 246
 DB 180 KIOPDVALPYSSGTGLPKGVMLTKHGLVTSVAQVDEGANLVHMSDVLKCVLDF 239
 QY 247 HGFQFETLIGLYTCGFR---IVMLTKRDEEFELKTDYKCSVITLYPTEFALINRSEL 302
 DB 240 H---IYSLNSVLLCALRGAAALLIMOKFDIAQLELIRPKHVTIGPEPPVPLAIAKSP 296
 QY 303 LDKYDLSNVEIASGAPLSEKIGEAVARREMLPGVROGYGLTETTSAILI-----TPEG 357
 DB 297 VINNYDLSVRVYMSGAPLSEKIGEAVARREMLPGVROGYGLTETTSAILI-----TPEG 356
 QY 358 DDKPGASGVPLERAKYIDLDTKKTLGPNRRGEVCKGPMIMKGYVNDPEATREIIDE 417
 DB 357 DKSAGCTGVVNAEMKIVDPDGTGSLPNRPGECICIRDDIMKGYLNDPEATREIIDE 416
 QY 418 GMLHNGDIGYDEKHEFFIVDLKSLIKKYGVPAALESVLLHNPINPDAGVGVDP 477
 DB 417 GMLHNGDIGYDEKHEFFIVDLKSLIKKYGVPAALESVLLHNPINPDAGVGVDP 476
 QY 478 IAGELPGAVVVLKKGKSMTEKEMDVVASOVSNARKLRGVRFVDEVRKGLGKIDGKAI 537
 DB 477 QNGEYVAVVAVNSNSTIREDVKDFISKQVIFYKRIK-RVFEVEVTPKSPSGKILRKDL 535

QY 538 R 538
 DB 536 R 536

RESULT 13
 ID 4CL2-SOYBN STANDARD; PRT; 562 AA.
 AC P1687;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE 4-COMARATE--COA LIGASE 2 (EC 6.2.1.12) (4CL 2) (4-COMAROYL-COA
 DE SYNTHASE 2) (CLONE 4CL16).
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lindermayr C.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 140-562 FROM N.A.
 RC STRAIN=CV. HAROSY 63;
 RX MEDLINE=94105342; PubMed=8278545;
 RA Uhlmann A., Ebel J.;
 RT "Molecular cloning and expression of 4-coumarate:coenzyme A ligase,
 RT an enzyme involved in the resistance response of soybean (Glycine max
 RT L.) against pathogen attack.";
 RL Plant Physiol. 102:1147-1156(1993).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COMARATE + COA = AMP + DIPHOSPHATE +
 CC 4-COMAROYL-COA.
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.

 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: X69955; CAC36095.1; -.
 DR PIR: S31709; S31709.
 DR PIR: P00772; P00772.
 DR HSSP; P08659; ILCI.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 1.
 DR PROSITE; PS00455; AMP_BINDING; 1.
 KW Ligase; Phenylpropanoid metabolism; Multigene family.
 SQ SEQUENCE 562 AA; 60963 MW; 2AB4652DDA5160B2 CRC64;

Query Match 25.7%; Score 726; DB 1; Length 562;
 Best Local Similarity 33.2%; Pred. No. 3.7e-43;
 Matches 187; Conservative 100; Mismatches 239; Indels 38; Gaps 11;

QY 1 MENNENDENIYGE-----PYPIEGSGAGOLKRYMDRYAKGAIATNAL--- 48
 DB 8 LDFPKTDQNGVSDPQTSVHFESKLPDIP---SNHPLHSYC--FQNLSPFAHRPCLTVG 62
 QY 49 -TGVDYTYAELEKSCCLGALKNYGLVVDGRIALGSENCBEFFIPVLAFLGCVGAP 107
 DB 63 PAKTFYADTHLISSTKIAAGLSMGLKGDVYVITLLQNSADFFSLAISMGAVATTA 122
 QY 108 NEIYTLRELHSLGICRPTVFSKGLDVIYVQ--KTYFAIKTIVILDSKYDGYGOS 165
 DB 123 NPFTAPETIRQPIVTSKAKLIITQAMTVVDKLRNHDGAKLGSDPFVVVVDPPENCLEHFSV 182

Db 473 LEALLIGHPDITDVAAMKEEAGEVPFAFVVKSKDSELSDDKQFVSKOVFYKRI- 531

Qy 516 GGVRFVDEVPKGLTGKIDGKAIR 538

Db 532 NKVFETESIPKAPSGKILRKDLR 554

RESULT 15

4CL2_ARATH STANDARD; PRT; 556 AA.

ID 4CL2_ARATH

AC 095725: 091035:

DT 20-AUG-2001 (Rel. 40, Created)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE 4-COMARATE-COA LIGASE 2 (EC 6.2.1.12) (4CL 2) (4-COMAROYL-COA SYNTHASE 2).

GN 4CL2 OR AT3G21240 OR MXL8.10.

OS Arabidopsis thaliana (mouse-ear cress).

OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eucosids II: Brassicales: Brassicaceae: Arabidopsis.

OC NCBI_TaxID=3702;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RX MEDLINE=9348176; PubMed=10417722;

RA Eutling J., Buetlner D., Wang Q., Douglas C.J., Somasich I.E., Kombrink E.;

RT "Three 4-coumarate:coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms.";

RL Plant J. 19:9-20(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RX MEDLINE=20277480; PubMed=10819329;

RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence features of the regions of 4,504,864 bp covered by sixty P1 and YAC clones.";

RL DNA Res. 7:131-135(2000).

CC -1- CATALYTIC ACTIVITY: ATP + 4-COMARATE + COA = AMP + DIPHOSPHATE + 4-COMAROYL-COA.

CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.

CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF106085; AAD47192.1; -

DR EMBL: AF106086; AAD47193.1; -

DR EMBL: AB023045; BAB01716.1; -

DR HSSP: P08659; 1LC1.

DR InterPro: IPR000873; AMP-bind.

DR Pfam: PF00501; AMP-binding; 1.

DR PROSITE: PS00455; AMP-BINDING; 1.

KW Ligase: Phenylpropanoid metabolism; Multigene family.

FT CONFLICT 247 W -> L (IN REF. 2).

FT CONFLICT 265 I -> V (IN REF. 2).

SEQUENCE 556 AA; 60929 MW; E3C2EAD677529EED CRC64;

Query Match 25.2%; Score 712; DB 1; Length 556;

Best Local Similarity 34.8%; Pred. No. 3.4e-42;

Matches 178; Conservative 91; Mismatches 204; Indels 38; Gaps 8;

Qy 46 NALTGVDTYAEYLEKSCCLGEALKNYGLVVGRIALCSNCEFFPIVLAGLIGVGA 105

Db 57 NGPTGEVYTTADVHTSRKLAAGLHNGKHNDVWMLLPNSPEVLLTFPLASFGAIRT 116

Qy 106 PINEIYTLRELVSLSISKPTIVFSSKGLDKVITVOK-----TVT-----AIRTIVLDSK 157

Db 117 SANDEFTPAEISKQAKASAKALIVTOSRYVDKIKNLQNDGVLIVTDSALPENCILRSE 176

Qy 158 VDRGYQSDMNFIKKNTPOGFKSGSFYEVNKRQVALIMNSGSGTGLPKGVOLTHENL 217

Db 177 LTQSEPRVDSIEKISP-----EDVALPFSGGTGLPKGVMLTHKGL 220

Qy 218 VTRFSHARD-----PIYGNQVSPETALITVPPFHGGMFT-TLGYITGCFRIVMLTKRDE 272

Db 221 VTSVAQVQDENNLTFNR-----DDVILCVMPHPIYALNSIMCSLRIGATLLIMPKFEI 277

Qy 273 EYFLKTLQDYKCSVLIVLTFALINRSELDDKYDLSNVEIASGAPLSKEIGAVARR 332

Db 278 TLLLEQIQCKKVAVMVPPIVLAIAKSPETEKEYDLSVBMWKSQAAPLGRLEDAISAK 337

Qy 333 FNLPGVROGVLGETTSALITTPEGDKP-----GASGVVPLFAKVIDLDTKKTLPGN 387

Db 338 FPNAKLGQGYGMEAGPVLAMSLGFAKEFPYKSGCGTVVNNAEKILDPDTGSLPRN 397

Qy 388 RRGECVCKGPMKMGVYDNPDEATREIIDEQWLHGTGIDGYDEEKNHFTYDLSKLITRK 447

Db 398 KPEGLCIRGNQIMKGYLNDPLATASTIDKQGLHGTGIDGFIIDDDLEFLVDRKLKELIKK 457

Qy 448 GYQVPAELLESVLLQPNIFDAGVAGVDPPIAGELPGAVVYLKKGSKMTEKEMDYVASQ 507

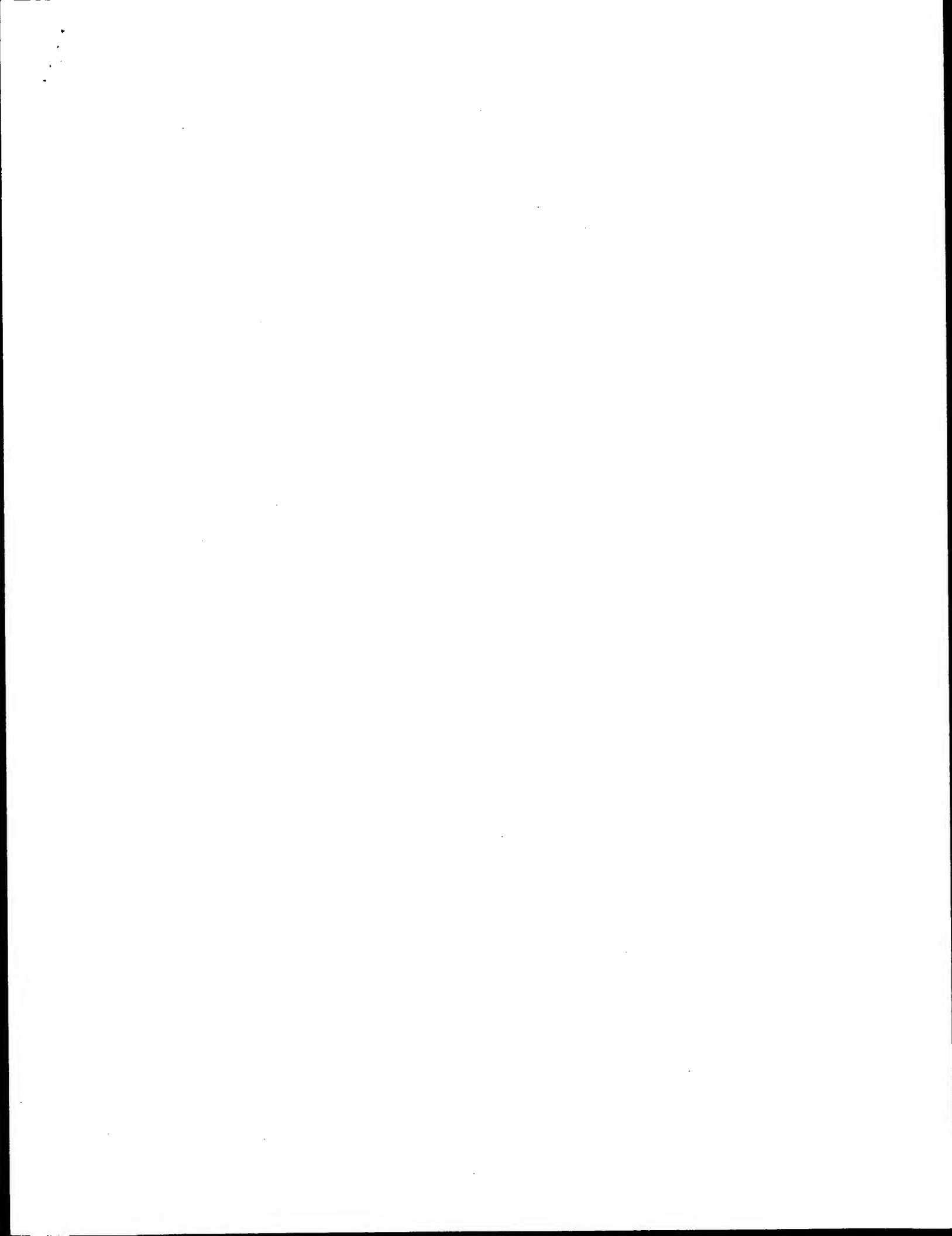
Db 458 GQVAPAELESLLIGHPEINDVAVAMKEDEAGEVPFAFVVKSKDSELSDDKQFVSKOVFYKRI 517

Qy 508 VSNAKRLRGVRFVDEVPKGLTGKIDGKAIR 538

Db 518 VVFYKRI-NKVFETESIPKAPSGKILRKDLR 547

Search completed: November 29, 2001, 02:00:07

Job time: 377 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 02:00:07 ; Search time 25.81 Seconds
(without alignments)
778.471 Million cell updates/sec

Title: US-09-581-241-6

Perfect score: 2823

Sequence: 1 MNNMNDENIVGPEPPYPI.....TGKIDGKAIKREILKKPKAKM 548

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2814	99.7	548	1	Q01158 luciola lat
2	2684	95.1	548	1	1UC1_LUCCLA
3	1945.5	68.9	550	1	1UC1_LUCGR
4	1766.5	27.2	544	1	1UC1_PETCR
5	1758.5	26.9	544	1	1UC1_PETCR
6	1756.5	26.8	547	1	1UC1_TOBAC
7	1746.5	26.4	542	1	1UC1_TOBAC
8	1745	26.4	553	1	1UC1_VANPL
9	1732.5	25.9	545	1	1UC1_SOLTU
10	1731	25.9	561	1	1UC1_ORYSA
11	1730	25.9	563	1	1UC1_ORYSA
12	1725.3	25.7	545	1	1UC1_SOLTU
13	1724	25.6	561	1	1UC1_ORYSA
14	1718.5	25.5	562	1	1UC1_ARYTH
15	1710	25.2	556	1	1UC1_ARYTH
16	1691.5	24.5	569	1	1UC1_ORYSA
17	1672	23.8	537	1	1UC1_PINFA
18	1587.5	20.8	293	1	1UC1_SOYBN
19	1574.5	20.4	560	1	1UC1_BACSU
20	1568	20.1	561	1	1UC1_BACSU
21	1565	17.1	562	1	1UC1_HAELN
22	1465	16.5	543	1	1P4T2_YEAST
23	1460	16.3	572	1	1ACSA_BACSU
24	1446	15.8	548	1	1YDID_ECOLI
25	1425	15.1	522	1	1CAIC_ECOLI
26	1370	13.1	698	1	1LCFB_HUMAN
27	1368	13.0	699	1	1LCFB_HUMAN
28	1361	12.8	699	1	1LCFB_MOUSE
29	1357	12.6	492	1	1MENE_STRAU
30	1353.5	12.5	699	1	1LCFA_HUMAN
31	1352	12.5	699	1	1LCFB_HUMAN
32	1348.5	12.3	693	1	1LCFE_HUMAN
33	1340	12.0	698	1	1LCFA_CAVPO

ALIGNMENTS

RESULT	ID	STANDARD	PRT	548 AA
1	LUC1_LUCCLA			
AC	001158:			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).			
OS	Luciola lateralis (firefly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;			
OC	Elateriformia; Cantharidea; Lampyridae; Luciola.			
OX	NCBI_TaxID=7052;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92305054; PubMed=1610896;			
RA	Tatsumi H., Kajiyama N., Nakano E.;			
RT	"Molecular cloning and expression in <i>Escherichia coli</i> of a cDNA clone			
RT	encoding luciferase of a firefly, <i>Luciola lateralis</i> .";			
RL	Biochim. Biophys. Acta 1131:161-165(1992).			
CC	-1- FUNCTION: PRODUCES GREEN LIGHT.			
CC	-1- CATALYTIC ACTIVITY: LUCIFERIN + O(2) + ATP -> OXIDIZED LUCIFERIN +			
CC	CO(2) + H(2)O + AMP + PYROPHOSPHATE + LIGHT.			
CC	-1- COFACTOR: REQUIRES MAGNESIUM.			
CC	-1- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME			
CC	FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL: X66919; CAA47358.1; -			
DR	PIR: S23437; S23437.			
DR	HSSP: P08659; 1UC1.			
DR	InterPro: IPR000873; AMP-bind.			
DR	Pfam: PF00501; AMP-binding.1.			
DR	PROSITE: PS00342; MICROBODIES_CTER: 1.			
DR	PROSITE: PS00455; AMP-BINDING: 1.			
KW	Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;			
KW	Peroxisome.			
FT	SITE 546 548 MICROBODY TARGETING SIGNAL (POTENTIAL).			
SEQ	SEQUENCE 548 AA: 60125 MW: AC62F9320BB6D4A6 CMC64;			
Query Match	99.7%; Score 2814; DB 1; Length 548;			
Best local similarity	99.6%; Pred. No. 2.8e-188;			
Matches 546:	Conservative 1; Mismatches 1; Indels 0; Gaps 0;			
QY	1 MNNMNDENIVGPEPPYPIIEGSGAGQLRKMYDRYAKGIAFTNALGVDTYATYLE 60			
DB	1 MNNMNDENIVGPEPPYPIIEGSGAGQLRKMYDRYAKGIAFTNALGVDTYATYLE 60			

```

QY 61 KSCCGEALKNYGLVYDGRIALCSENCEEFFIPVLAGLFIGVGAFTNEIYTLRELHSL 120
DB 61 KSCCGEALKNYGLVYDGRIALCSENCEEFFIPVLAGLFIGVGAFTNEIYTLRELHSL 120
QY 121 GISKRTIYSSSKKGLDKVITVOKTATATITVILDSKVDYRGYSMDNFINKNTPOGFGK 180
DB 121 GISKRTIYSSSKKGLDKVITVOKTATATITVILDSKVDYRGYSMDNFINKNTPOGFGK 180
QY 181 SSFKTVEVNRKQVALIMNSSGSTGLPKGVOLTHENIYTRFSHARDPIYGNVSPGTAIL 240
DB 181 SSFKTVEVNRKQVALIMNSSGSTGLPKGVOLTHENIYTRFSHARDPIYGNVSPGTAIL 240
QY 241 TVPPEHGFMTTLLGYLTCGFRIVMLTKFDETEFLKTLQDYKCSSVILVPLFALINRS 300
DB 241 TVPPEHGFMTTLLGYLTCGFRIVMLTKFDETEFLKTLQDYKCSSVILVPLFALINRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAARFNLPGVROGGLTETTSAILITTEBGDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAARFNLPGVROGGLTETTSAILITTEBGDK 360
QY 361 PGASGVVPLFKAKYIDIDTKKTLGPNRRGEVCVKGPMIMKGYVNPPEATREIIDEBCWL 420
DB 361 PGASGVVPLFKAKYIDIDTKKTLGPNRRGEVCVKGPMIMKGYVNPPEATREIIDEBCWL 420
QY 421 HTGDIQYDEEKNHFTVDRSLIKYKGQVPPAELESVLLQHPNIPDAGVAGVDPPIAG 480
DB 421 HTGDIQYDEEKNHFTVDRSLIKYKGQVPPAELESVLLQHPNIPDAGVAGVDPPIAG 480
QY 481 ELLPGVAVVLLKKGKSMTEKEVMDYVASOVSNKRLRGVRFVDEVKGLTKIDGKAIRFI 540
DB 481 ELLPGVAVVLLKKGKSMTEKEVMDYVASOVSNKRLRGVRFVDEVKGLTKIDGKAIRFI 540
QY 541 LKKPVAKM 548
DB 541 LKKPVAKM 548

RESULT 2
LUC1_LUCCR STANDARD; PRT: 548 AA.
ID LUC1_LUCCR STANDARD; PRT: 548 AA.
AC P13129;
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).
OS Luciola cruciata (Japanese firefly) (Genji firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharoidea; Lampyridae; Luciola.
OX NCBI_TaxID=7051;
RN [1]
RM SEQUENCE FROM N.A.
RX MEDLINE=69326143; PubMed=2473944;
RA Masuda T., Tatsuji H., Nakano E.;
RT "Cloning and sequence analysis of cDNA for luciferase of a Japanese
RL firefly, Luciola cruciata."
CC -1- FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 544 NM.
CC -1- CATALYTIC ACTIVITY: LUCIFERIN + O(2) + ATP = OXIDIZED LUCIFERIN +
CC CO(2) + H(2)O + AMP + PYROPHOSPHATE + LIGHT.
CC -1- COFACTOR: REQUIRES MAGNESIUM.
CC -1- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
CC EMBL: M26194; AAA29135.1;
CC DR PIR: J50181;
CC DR HSP: P08659; TICI.
CC DR InterPro: IPR000873; AMP-bind.
CC DR Pfam: PF00501; AMP-binding; 1.
CC DR ProSite: PS00342; MICROBODIES_CTER; 1.
CC DR ProSite: PS00455; AMP BINDING; 1.
CC KN Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
CC FT SITE 546 548 MICROBODY TARGETING SIGNAL (POTENTIAL).
CC SQ SEQUENCE 548 AA: 60017 MW: 2052D6189E79109F CRC64.

Query Match          95.1%; Score 2684; DB 1; Length 548;
Best Local Similarity 93.4%; Pred. NO. 3e-179;
Matches 512; Conservative 26; Mismatches 10; Indels 0; Gaps 0;

QY 1 MENMENDENIVYGPPEFYPIEBSAGALRKYMRYAKIGATFATNALGVDTYAYEYLE 60
DB 1 MENMENDENIVYGPPEFYPIEBSAGALRKYMRYAKIGATFATNALGVDTYAYEYLE 60
QY 61 KSCCGEALKNYGLVYDGRIALCSENCEEFFIPVLAGLFIGVGAFTNEIYTLRELHSL 120
DB 61 KSCCGEALKNYGLVYDGRIALCSENCEEFFIPVLAGLFIGVGAFTNEIYTLRELHSL 120
QY 121 GISKRTIYSSSKKGLDKVITVOKTATATITVILDSKVDYRGYSMDNFINKNTPOGFGK 180
DB 121 GISKRTIYSSSKKGLDKVITVOKTATATITVILDSKVDYRGYSMDNFINKNTPOGFGK 180
QY 181 SSFKTVEVNRKQVALIMNSSGSTGLPKGVOLTHENIYTRFSHARDPIYGNVSPGTAIL 240
DB 181 SSFKTVEVNRKQVALIMNSSGSTGLPKGVOLTHENIYTRFSHARDPIYGNVSPGTAIL 240
QY 241 TVPPEHGFMTTLLGYLTCGFRIVMLTKFDETEFLKTLQDYKCSSVILVPLFALINRS 300
DB 241 TVPPEHGFMTTLLGYLTCGFRIVMLTKFDETEFLKTLQDYKCSSVILVPLFALINRS 300
QY 301 ELLDKYDLSNLVEIASGAPLSKEIGEAARFNLPGVROGGLTETTSAILITTEBGDK 360
DB 301 ELLDKYDLSNLVEIASGAPLSKEIGEAARFNLPGVROGGLTETTSAILITTEBGDK 360
QY 361 PGASGVVPLFKAKYIDIDTKKTLGPNRRGEVCVKGPMIMKGYVNPPEATREIIDEBCWL 420
DB 361 PGASGVVPLFKAKYIDIDTKKTLGPNRRGEVCVKGPMIMKGYVNPPEATREIIDEBCWL 420
QY 421 HTGDIQYDEEKNHFTVDRSLIKYKGQVPPAELESVLLQHPNIPDAGVAGVDPPIAG 480
DB 421 HTGDIQYDEEKNHFTVDRSLIKYKGQVPPAELESVLLQHPNIPDAGVAGVDPPIAG 480
QY 481 ELLPGVAVVLLKKGKSMTEKEVMDYVASOVSNKRLRGVRFVDEVKGLTKIDGKAIRFI 540
DB 481 ELLPGVAVVLLKKGKSMTEKEVMDYVASOVSNKRLRGVRFVDEVKGLTKIDGKAIRFI 540
QY 541 LKKPVAKM 548
DB 541 LKKPVAKM 548

RESULT 3
LUC1_PROPY STANDARD; PRT: 550 AA.
ID LUC1_PROPY STANDARD; PRT: 550 AA.
AC P08659;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LUCIFERIN 4-MONOOXYGENASE (EC 1.13.12.7) (LUCIFERASE).
OS Photinus pyralis (North American firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharoidea; Lampyridae; Photinus.
OX NCBI_TaxID=7054;
RN [1]

```

RA de Wet J.R., Wood K.V., Deluca M., Helinski D.R., Subramani S.;
 "Firefly Luciferase gene: structure and expression in mammalian
 cells."; Mol. Cell. Biol. 7:725-737(1987).
 [21]
 RN
 RP SUBCELLULAR LOCATION.
 RX MEDLINE-8720417; PubMed-3554235;
 RA Keller G.-A., Gould S., de Luca M., Subramani S.;
 "Firefly Luciferase is targeted to peroxisomes in mammalian cells."; Proc. Natl. Acad. Sci. U.S.A. 84:3264-3268(1987).
 RL
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE-96398615; PubMed-8805533;
 RA Conti E., Franks N.P., Brick P.;
 "Crystal structure of firefly luciferase throws light on a
 superfamily of adenylate-forming enzymes."; Structure 4:287-298(1996).
 RL
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE-99007339; PubMed-9788915;
 RA Franks N.P., Jenkins A., Conti E., Lieb W.R., Brick P.;
 "Structural basis for the inhibition of firefly luciferase by a
 general anesthetic."; Biophys. J. 75:2205-2211(1998).
 RL
 CC -1- FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 562 NM.
 CC -1- CATALYTIC ACTIVITY: LUCIFERIN + O(2) + ATP = OXIDIZED LUCIFERIN +
 CC CO(2) + H(2)O + AMP + PYROPHOSPHATE + LIGHT.
 CC -1- COFACTOR: REQUIRES MAGNESIUM.
 CC -1- SUBCELLULAR LOCATION: PEROXISOML.
 CC
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/announce/
 CC or_send_an_email_to_license@isb-sib.ch](http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch)).
 CC
 CC
 CC EMBL: M15077; AAA29795.1; -;
 CC EMBL: X84848; CA59283.1; -;
 CC EMBL: U03687; AAA03561.1; -;
 CC EMBL: U08934; AAB64396.1; -;
 CC EMBL: U89935; AAB64399.1; -;
 CC PIR: A26772; A26772.
 CC PDB: 1LCI; 26-MAR-97.
 CC PDB: 1BA3; 11-NOV-98.
 CC InterPro: IPR000873; AMP-bind.
 CC Pfam: PF00501; AMP-binding; 1.
 CC PROSITE: PS00342; MICROBODIES_CTER; 1.
 CC PROSITE: PS00455; AMP BINDING; 1.
 CC Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
 CC Peroxisome; 3D-structure.
 KW SITE 548 550 MICROBODY TARGETING SIGNAL.
 SQ SEQUENCE 550 AA; 60745 MW; E380FCE9056ACD6 CRC64;

Db	121	SQPIVVEVSKSGLOKILNVQKLPITIQKIIIMOSKTDYQGGQSMYTFVSHLPGENEYD	180
Qy	183	FKTVEVNRKEGOVALIMNMSGSTGLPKGQVOLTHERNIYRFSHARDPIYGNVSPGTALIV	242
Db	181	FVPESEFDKDKTIALIMNMSGSTGLPKGVALPHRACVRFSHARDPIFGNOIIPDTALSV	240
Qy	243	VPEHHGEGMPTTIGYLTGCGFRIVMLTFRDEETFLTLDYKCCSSVILVPTLALINSEL	302
Db	241	VPEHHGEGMPTTIGYLTGCGFRIVMLYRFEELFLRSIDQYKIGALLVPTLSEPFKASTL	300
Qy	303	LDKIDLSLVLIASGGAPLSEIEGAVARRNPLGVQNGYGLTFTTSAIIITPEBGDKPG	362
Db	301	IDKYDLSMLHBIASGGAPLSEKEVEBAARKEPHLGIQGGYGLTFTTSAIITPEBGDKPG	360
Qy	363	ASGVKVPFLFKKAVIIDLPKKTILGNRGEVYVCKPMLKMGVNDPEATREIIDEEGMLHT	422
Db	361	AVGKAVPFEFEKAVVDIDGKTILGNQKRELCVCKPMLMSGYVNNPEATNALIDKCMHLS	420
Qy	423	GDIGYDEEKAFFIYDRLSLTIKKGYQVPAELESYLLQHPNIFDAGVADPPIAGEL	482
Db	421	GDIAWDEDEHFIYDRILKSLTIKKGYVANAELLESILLQHPNIFDAGVAGLPIDDAAGEL	480
Qy	483	PGAVVVLKKGSMTKEVEYMDVASQVSNAKRLRGVVRVDEVPKGLTGKIDGKAIREILK	542
Db	481	PAAVVVLHGHTKTEKEIYDVVASQVTTAKRLRGGVVFEVPEKGLTGKIDARKIREILI	540
Qy	543	K 543	
Db	541	K 541	

[illegible]

CC	RESULT	4
AC	ACL1_PETCR	STANDARD; PRT; 544 AA.
AC	P14912;	
DT	01-APR-1990	(Rel. 14, Created)
DT	01-APR-1990	(Rel. 14, Last sequence update)
DT	20-AUG-2001	(Rel. 40, Last annotation update)
DE	4-COMMARATE--COA LIGASE 1 (EC 6.2.1.12) (4CL1) (4-COMMAROYL-COA SYNTHASE 1).	
DE	GN	4CL1 OR 4CL-1.
OS	Petroselinum crispum (Parsley) (petroselinum hortense).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.	
OC	NCBI_TaxID=4043;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=89005119; Pubmed=3169018;	
RA	Lozoya E., Hoffmann H., Douglas C., Schulz W., Scheel D., Halbrock K.;	
RT	"Primary structures and catalytic properties of isoenzymes encoded by the two 4-coumarate:CoA ligase genes in parsley.";	
RL	Eur. J. Biochem. 176:661-667(1988). [2]	
RN	SEQUENCE OF 1-8 FROM N.A.	
RP	Douglas C., Hoffmann H., Schulz W., Halbrock K.;	
RT	"Structure and elicitor or U.V.-light-stimulated expression of two 4-coumarate:CoA ligase genes in parsley.";	
RL	EMBO J. 6:1189-1195(1987).	
CC	-1- CATALYTIC ACTIVITY: ATP + 4-COMMARATE + COA = AMP + DIPHOSPHATE + 4-COMMAROYL-COA.	
CC	-1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.	
CC	-1- INDUCTION: TRANSCRIPTIONALLY ACTIVATED BY BOTH FUNGAL ELICTOR AND ULTRAVIOLET IRRADIATION.	
CC	-1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its	

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X13324; CAA31696.1; -
DR EMBL: X05350; CAA28959.1; -
DR PIR: S01667; S01667.
DR HSSP: P08659; 1ICI.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00453; AMP_BINDING; 1.
DR Ligase: Phenylpropanoid metabolism; Multigene family.
KW SEQUENCE 544 AA; 59825 MW; 22BBAD78F255D0C8 CRC64;

Query Match 27.2%; Score 766.5; DB 1; Length 544;
Best Local Similarity 36.0%; Pred. No. 6.3e-46;
Matches 186; Conservative 89; Mismatches 189; Indels 53; Gaps 10;

QY 46 NALTGVYVYAEYLEKSCGLGALKNYGLVNDGRIALCSENEEFFPIVLAFLIGVYA 105
DB 47 NCATGETFYVSGVELLSKRVASGLNKLGIQGDITMLLPNSPEYFFALGASYRGAIST 106
QY 106 PNEIYVTLRELHSLGSKPTIVFSSKGLDRTVYQKTVTAIKTIVIL--DSKVDYRG 162
DB 107 MANPFTSAEVIKOLKASQAKLITQACVYDKV---KDYAEKNQIICIDDAPODCH 162
QY 163 YQSMNFIKKNTPOGFGKSSFTVEYNRKEQVALIMNSSGSGTGLPKGVQLTHENTVTR-- 220
DB 163 FSKLMEADSEMEP-----VVINSDDVAL--PYSSGTTGLPKGVMLTKGLVTSVA 212
QY 221 -----FSHARDPIYGNQVSPGTALIVPPHHGGMFTTGLYLCGR---IYM 266
DB 213 QQVGDNDPNLYMHSB-----VMICILPLFH---IYSLNVLCCGLRAGVTIILI 258
QY 267 LTRPDEETFLKTLQDYKCSSVILVPTLFAILNRSELLDKYDLSNVEIASGAPLSKETG 326
DB 259 MOKFPIVPELELIQKYKTVIGFVPPYIALAKSPVVDYDLSSTVYVWSGAPLKELE 318
QY 327 EAVARFNLPGVRGQYGLTETTSATII-----TPGDDKPGASGVVPLFKAKVIDLDTK 381
DB 319 DAVRAKFPNAKIGQGYGMEAGPVLAACLAFAKEPEYIKSGAGTVYVNAEMKIYDPTN 378
QY 382 KTLGNRRGECVCKPMLKGYVNDPEATRELIIDEGWLHTDIDIGYDEKHEFFIVDRK 441
DB 379 ASLPRNKGELICIRGDQIMKGYLNDPESTRITIDEGWMLHTDIDGIDDELFTVDRK 438
QY 442 SLIRKGYVPAELESVILQHPNIFDAGVAGVPPPIAGELPGAVVYLKKGSMTEKEVM 501
DB 439 ELIKKKGQVAAEELALLHTPTSDAIVPMIDEKAGEVAVAVRTNGFTTEEEK 498
QY 502 DYVASOVNAKRLRGVRVDEVPKGLTGKIDGKAIR 538
DB 499 QFVSKQVVFYKRI-FRVFFVDAIIPKSPSGKILRKDR 534

RESULT 5
ACCL2_PETCR STANDARD: PRT: 544 AA.
ID ACCL2_PETCR
AC P14913;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 4-COUMARATE--COA LIGASE 1 (EC 6.2.1.12) (ACCL 1) (4-COUMAROYL-COA
DE SYNTHASE 1).
GN 4CL2 OR 4CL 2.
OS Petroselinum crispum (Parsley) (Petroselinum hortense).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
OC NCBI_TaxID=4043;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE-89005119; PubMed-3169018;
RA Lozoya E., Hoffmann H., Douglas C., Schulz W., Scheel D.,
RA Hahlbrock K.,
RT "Primary structures and catalytic properties of isoenzymes encoded by
RT the two 4-coumarate:COA ligase genes in parsley."
RN Eur. J. Biochem. 176:661-667(1988).
RN [2]
RP SEQUENCE OF 1-8 FROM N.A.
RA Douglas C., Hoffmann H., Schulz W., Hahlbrock K.;
RT "Structure and elicitor or U.V.-light-stimulated expression of two
RT 4-coumarate:COA ligase genes in parsley."
RL EMBO J. 6:1189-1195(1987).
CC -1- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
CC 4-COUMAROYL-COA.
CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -1- INDUCTION: TRANSCRIPTIONALLY ACTIVATED BY BOTH FUNGAL ELICITOR AND
CC ULTRAVIOLET IRRADIATION.
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X13325; CAA31697.1; -
DR EMBL: X05351; CAA28960.1; -
DR PIR: S01667; S01667.
DR PIR: S15695; S15695.
DR HSSP: P08659; 1ICI.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00453; AMP_BINDING; 1.
KW Ligase: Phenylpropanoid metabolism; Multigene family.
KW SEQUENCE 544 AA; 59783 MW; B477965C68F8C534 CRC64;

Query Match 26.9%; Score 758.5; DB 1; Length 544;
Best Local Similarity 35.8%; Pred. No. 2.3e-45;
Matches 185; Conservative 89; Mismatches 190; Indels 53; Gaps 10;

QY 46 NALTGVYVYAEYLEKSCGLGALKNYGLVNDGRIALCSENEEFFPIVLAFLIGVYA 105
DB 47 NCATGETFYVSGVELLSKRVASGLNKLGIQGDITMLLPNSPEYFFALGASYRGAIST 106
QY 106 PNEIYVTLRELHSLGSKPTIVFSSKGLDRTVYQKTVTAIKTIVIL--DSKVDYRG 162
DB 107 MANPFTSAEVIKOLKASQAKLITQACVYDKV---KDYAEKNQIICIDDAPODCH 162
QY 163 YQSMNFIKKNTPOGFGKSSFTVEYNRKEQVALIMNSSGSGTGLPKGVQLTHENTVTR-- 220
DB 163 FSKLMEADSEMEP-----VVINSDDVAL--PYSSGTTGLPKGVMLTKGLVTSVA 212
QY 221 -----FSHARDPIYGNQVSPGTALIVPPHHGGMFTTGLYLCGR---IYM 266
DB 213 QQVGDNDPNLYMHSB-----VMICILPLFH---IYSLNVLCCGLRAGVTIILI 258
QY 267 LTRPDEETFLKTLQDYKCSSVILVPTLFAILNRSELLDKYDLSNVEIASGAPLSKETG 326
DB 259 MOKFPIVPELELIQKYKTVIGFVPPYIALAKSPVVDYDLSSTVYVWSGAPLKELE 318
QY 327 EAVARFNLPGVRGQYGLTETTSATII-----TPGDDKPGASGVVPLFKAKVIDLDTK 381
DB 319 DAVRAKFPNAKIGQGYGMEAGPVLAACLAFAKEPEYIKSGAGTVYVNAEMKIYDPTN 378
QY 382 KTLGNRRGECVCKPMLKGYVNDPEATRELIIDEGWLHTDIDIGYDEKHEFFIVDRK 441
DB 379 ASLPRNKGELICIRGDQIMKGYLNDPESTRITIDEGWMLHTDIDGIDDELFTVDRK 438

QY 442 SLIKYGYVPAELESVLLQHPNIFDAGVAGVDPPIAGELPGAVVVLKKGSMTEKEYW 501
 DB 439 ELIKYKGFVPAELEMALTLTHTPTISDAVAVPIDERAGVPAFVAVRTNGFTTEELK 498
 QY 502 DYVASQVSNARKRLRGVREVEDEPKGLTGKIDGKAIR 538
 DB 499 QFVSKGVVFFYKRI-FRFFVFDALPKSPSKILKKDLR 534

RESULT 6
 4CL1_TOBAC STANDARD; PRT: 547 AA.
 AC 024145;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 4-COUMARATE--COA LIGASE 1 (EC 6.2.1.12) (4CL 1) (4-COUMAROYL-COA
 DE SYNTHASE 1).
 GN 4CL1.
 OS Nicotiana tabacum (Common tobacco)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OC NCBI_TaxID=4097;
 RX MEDLINE=96416441; PubMed=8819324;
 RA Lee D., Douglas C.J.;
 RT "Two divergent members of a tobacco 4-coumarate:coenzyme A ligase
 RT (4CL) gene family. cDNA structure, gene inheritance and expression,
 RT and properties of recombinant proteins."
 RL Plant Physiol. 112:193-205(1996).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
 CC 4-COUMAROYL-COA.
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U50845; AAB18637.1; -
 CC HSSP: P08659; ILCI.
 CC Mendel: 24477; Nicita:1179;24477.
 CC InterPro: IPR000873; AMP-bind.
 CC Pfam: PF00501; AMP-binding; 1.
 CC PROSITE: PS00455; AMP-BINDING; 1.
 CC LIGASE: Phenylpropanoid metabolism; Multi-gene family.
 CC SEQUENCE 547 AA: 59842 MW: 154DF6D684E3F51B CRC64;
 SO

Query Match 26.88; Score 756.5; DB 1; Length 547;
 Best Local Similarity 34.98; Pred. No. 3.2e-45;
 Matches 188; Conservative 90; Mismatches 205; Indels 55; Gaps 12;

QY 32 YMDRRAKGAIAFNN-----ALNGVD--YTVAEVLEKSC-CGEMALKNYGLVVDG 78
 DB 25 YIPKHLPHISCFFENISFSSRPLINGANDQITVAF-VELTRKRAVGLNKIGIOKD 83
 QY 79 RIALSCENCEEFFIPLVAGLIGVAVPTMEITYTLRELVHSLGSKPTIVFSSKKGLDKV 138
 DB 84 TIMILLPSPPEFVAFMGASYLGAISTMANPLFPFAEVVQAKASAKIITQSCFVGKV 143
 QY 139 ITVQKTVAIKTIYILDSKVYRGYSMDNFIKNTROG-----FKSGSKVEVAV-R 190
 DB 144 -----KDYASENDVKYICIDSAPEGCIHFSELTQSDHEHIEPVKIQ 184

QY 191 KEOVALIMNSGSGTGLPGVOLTENITVRESHARDPIYGN-QVSPGTAILLVVPFHGF 249
 DB 185 PDDVALPYPSSGTGLPGKVMILTHKGLVTSVQOVDGENANILYMSEVYLMCVLPFH-- 242
 QY 250 GMEFTLGLYTCGFR-----IVMLTKFDEETFLKTLQDYACSSVILPTLAILNRSELDK 305
 DB 243 -IYLSNLSILGLRGAAILIMQKFDIAFPLELLIGKYVSGISGPEFVPLVLAIAKSPYDS 301
 QY 306 YDLSNLVETASGAPLSKEIGAVARRENPGVROGQITFTTSAILI-----IPEGDK 360
 DB 302 YDLSVYRTWSGAAPLGKLEDAVTRKFPNAKLGGYGTEGAPVLAACLAFAKPEFDIK 361
 QY 361 PGASGVVPLFAKAVYIDLTPTKTLGNRRGECVAGPMLMGVYDNPATREIIDEGML 420
 DB 362 SGACGIVVNAEMKIVDPPTGCSLPRNGELICINGDQIMKCYLNDPATRTIDKEGML 421
 QY 421 HTGDIGYDEEKHFIVDLRLKSLIRYKGYVPAELESVLLQHPNIFDAGVAGVDPPIAG 480
 DB 422 HTGDIGFIDEDELFIYDLRLKELIKYKGFVPAELEMALTLNHPNISDAVAVPMKDEQAG 481
 QY 481 ELPGAVVVLKKGSKMTEKEYMDYVASQVSNARKRLRGVREVEDEPKGLTGKIDGKAIR 538
 DB 482 EYPVAVFVRSNGSAITDEVKDFISKQVLYFYKRVK-RVFEFTVPSKSPSKILRRDLR 538

RESULT 7
 4CL2_TOBAC STANDARD; PRT: 542 AA.
 AC 024146;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 4-COUMARATE--COA LIGASE 2 (EC 6.2.1.12) (4CL 2) (4-COUMAROYL-COA
 DE SYNTHASE 2).
 GN 4CL2.
 OS Nicotiana tabacum (Common tobacco)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OC NCBI_TaxID=4097;
 RX MEDLINE=96416441; PubMed=8819324;
 RA Lee D., Douglas C.J.;
 RT "Two divergent members of a tobacco 4-coumarate:coenzyme A ligase
 RT (4CL) gene family. cDNA structure, gene inheritance and expression,
 RT and properties of recombinant proteins."
 RL Plant Physiol. 112:193-205(1996).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
 CC 4-COUMAROYL-COA.
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U50846; AAB18638.1; -
 CC HSSP: P08659; IBA3.
 CC Mendel: 24478; Nicita:1179;24478.
 CC InterPro: IPR000873; AMP-bind.
 CC Pfam: PF00501; AMP-binding; 1.
 CC PROSITE: PS00455; AMP-BINDING; 1.
 CC LIGASE: Phenylpropanoid metabolism; Multi-gene family.
 CC SEQUENCE 542 AA: 59479 MW: CB5579AEDFCF003 CRC64;
 SO


```

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91217100; PubMed=2022667;
RT Becker-Andre M., Schulze-Isferl P., Hahlbrock K.;
RT "Structural comparison, modes of expression, and putative cis-acting
RT elements of the two 4-coumarate: CoA ligase genes in potato.";
RL J. Biol. Chem. 266:8551-8559(1991).
CC -1- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
CC 4-COUMAROYL-COA.
CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M62755; AAA33842.1; -
CC DR PIR; A39827; A39827.
CC DR HSSP; P08659; ILCT.
CC DR InterPro: IPR000873; AMP-bind.
CC DR Pfam: PF00501; AMP-binding; 1.
CC DR PROSITE; PS00455; AMP-BINDING; 1.
CC KW Ligase; Phenylpropanoid metabolism; Multigene family.
CC SQ SEQUENCE 545 AA; 59619 MW; DE183683B74BA71 CRC64;

Query Match      25.9%; Score 732.5; DB 1; Length 545;
Best Local Similarity 34.0%; Pred. No. 1.5e-43;
Matches 184; Conservative 94; Mismatches 202; Indels 61; Gaps 12;

OY 32 YMDRYAKLAIFFTN-----ALTGVD--VTYAEYLEKSCCLGEALKNYGVNDR 79
DB 23 YIPKHLPLSYCFENSEFNSRPLCDGANDITTYAEVELTSKRVANGLKLTGQOKDT 82
OY 80 TALSCNCEFFTPVLAFLFVGVAAPNEITYELRELVHSLGDISKPTVFSKRGIDRYI 139
DB 83 IMLLEPCPEFVFAFAGSYLGAISTMANPLTPPAVVOQAKASAKIYITQACFAGKV- 141
OY 140 TVQKVTATKTIYIIDSQVDRYGYSKDNFIR--KNTPQG-----FKSSKTYEV 188
DB 142 -----KDY---AIEDMLKVICVDSPEGCVHFSELIQSDHEHPIFY 179
OY 189 N-RKEQVALINMSGSTGLPKGVOLTHENIVYRFSHARDPIYGN-QVSPGTAILLVVPFH 246
DB 180 KIQPPDVVALPSSSGTGLPKVMTLTHKGLVTSVAQVDGEMANLYMSDDYLCVLPDLF 239
OY 247 HGFQMTTLYLTGFR---IVMLTKRDEETFLTKLDYKCSSVILVPTFAILNRSSEL 302
DB 240 H---IYSLNSVYLCAIRYGAAILIMQKFDIAQFLLEIKHKVITLPPRYVLAIAKSP 296
OY 303 LKQYLSNVELJASGAPLSKIEGAVARFPLPGVRGQYGLTETTSALIT-----TPEG 357
DB 297 VNNYDSSVRYWASGAPLGELESAVRAKPRNAKLGQYGTGTEGAPVLAICLAFAPKEPF 356
OY 358 DDKPGASGVVPLDFKAKYIDLTGKTKLGNRRGECVAGKPMKGYNDNPETREIIDE 417
DB 357 DIKSGACGVVNAEMKLYNDPTGSLPRNGDELICINGDDIMKGYLNDPENTATILKE 416
OY 418 GMLHNGDIDGYDEKHFVYDLAKSLIYKGYVPALESVLLQHPNIFDAGVAGVDP 477
DB 417 GMLHNGDIDFIDDELFLYVRLKELIYKGFQVAPAELEALLIHNHPISDAAVVPIME 476
OY 478 IAGELPGAVVVLKKGKSMTEKEMVDYVAQVSNARKLGGVRFVDEYVKGILGRIDKAI 537
DB 478 IAGELPGAVVVLKKGKSMTEKEMVDYVAQVSNARKLGGVRFVDEYVKGILGRIDKAI 537

```

```

DB 477 QAGEVPAFVAVNSGTTTEDEVDKFIKQVIFYKRIK-RVFEVETVPKSPSGKILRKDL 535
OY 538 R 538
DB 536 R 536

RESULT 10
ACL3_ARATH
ID 4CL3_ARATH STANDARD; PRT; 561 AA.
AC Q95777;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 4-COUMARATE-COA LIGASE 3 (EC 6.2.1.12) (4CL 3) (4-COUMAROYL-COA
DE SYNTHASE 3).
GN 4CL3 OR ATL65060 OR F16316.6.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=99348176; PubMed=10417722;
RA Ehrling J., Buettner D., Wang Q., Douglas C.J., Somssich I.E.,
RA Kombirink E.;
RT "Three 4-coumarate:coenzyme A ligases in Arabidopsis thaliana
RT represent two evolutionarily divergent classes in angiosperms.";
RL Plant J. 19:9-20(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.-J., Koo H.-L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Laing-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marzalli A.,
RA Miltschker J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
RN [3]
RP CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
RP 4-COUMAROYL-COA.
CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF106087; AAD47194.1; -
CC DR EMBL; AF106088; AAD47195.1; -
CC DR EMBL; AC009360; AAF06039.1; -
CC DR HSSP; P08659; ILCT.

```


Db 473 LEALLIGHTPDTTDAVVAAMKEEAGEVPALFVYVSKDSSEDDVKQFVSQKVFFYKRI- 531

Oy 516 GGRFDEVEKGLTKIDKAIR 538
| : : : : :
Db 532 NKVFETESIPKPSGKILRKDLR 554

RESULT 15
ACCL2.ARAATH

ID	ACCL2.ARAATH	STANDARD;	PRT;	556 AA.
AC	09S725; 09LU035;			
DT	20-AUG-2001 (Rel. 40, Created)			
DT	20-AUG-2001 (Rel. 40, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	4-COMMATE--COA LIGASE 2 (EC 6.2.1.12) (ACL 2) (4-COMMAYOL-COA SYNTHASE 2)			
GN	ACCL2 OR AT13G21240 OR MXL8.10.			
OS	Arididopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.			
RN	NCBI_TaxId=3702;			
RM	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=CV. COLUMBIA;			
RX	MEDLINE=99348176; PubMed=10417722;			
RA	Hilting J., Buetner D., Wang Q., Douglas C.J., Somssich I.E., Kombrink E.;			
RT	"Three 4-commate:coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms."			
RL	Plant J. 19:9-20(1999).			
RP	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. COLUMBIA;			
RX	MEDLINE=20277480; PubMed=10819329;			
RA	Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;			
RT	"Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence features of the regions of 4,504,864 bp covered by sixty pl and TAC clones."			
RL	DNA Res. 7:131-135(2000).			
-I-	CATALYTIC ACTIVITY: ATP + 4-COMMATE + COA = AMP + DIPHOSPHATE + 4-COMMAYOL-COA.			
-I-	PATWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.			
-I-	SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF106085; AAD47192.1; -			
DR	EMBL; AF106086; AAD47193.1; -			
DR	EMBL; AB023045; BAB01716.1; -			
DR	HSSP; P08659; ILCI			
DR	InterPro; IPRO00873; AMP-bind.			
DR	Pfam; PF00501; AMP-binding.1.			
DR	PROSITE; PS00455; AMP_BINDING.1.			
KW	Ligase; Phenylpropanoid metabolism; Multigene family.			
FT	CONFLICT 247 247 W -> L (IN REF. 2).			
FT	CONFLICT 265 265 I -> V (IN REF. 2).			
SQ	SEQUENCE 556 AA; 60929 MW; EC3ZBAD77529EBD CRC64;			

```

Db      57 NCPTGVVYIYADVHTSRKLAAGLHNLGCKQDVMILLPNSPEVYVLFNLASFIGATTT 111
OY      106 PPHFEIYTLRELVHSGISKPTFVSSKGLDKVITYVK----TVI----AITYIYLDK 15
Db      117 SANPFTTPEISKOAKASAKLITVOSKRYVDIKRLQNDGVLYITVDSDAIPENCRLRESE 176
OY      158 VDYKGTQSDMNFIRKNTPDGFGKSSFKTVEVNRKQVALLMNSSGTGLPKGVOTHEHI 21
Db      177 LTQSEBPRVDISPEKISP-----EDVALPFSSTGTLPGVMYLTHTHGL 22
OY      218 VYRFSHAR-----PYGNOVSPGATLLVVPVPHHFGAFT--TLGYLTGCFYIMLYTFDE 27
Db      221 VYSAVOQVDGENPILYFN--DVIYLCWPFHFIYALNSIMLSLRIGATITLIMPREI 27
OY      273 EYFFLKTLDQKSSVILVPTFALLINRSLELDKYDLSNVEIASGAPLSEIGEVAAR 33
Db      278 TLLLEOTIRCKYAVMVPPIVLAIAKSPETEKEYDLSVRYMKSGAAPLKELEDAISAK 33
OY      333 FNLPGVROGYGLTETTSAILITPEBDDPR-----GASKVYPLRKAVYIDLDTKKTGPN 38
Db      338 FPNAKTLQGYGTEGVPYLAWSLGFAPKEPPRYKSGACSTVYRNAEMKILDDTGDSDLPN 39
OY      368 RRGVYCVAGPMILMKGYVUNPDATEIIDEBSWLTGDIYGYDEEKHFFIYDRKLSLTKY 44
Db      368 KPEGLCJINGNIMGYLMDPLATLASTIDKDMWLTGTVGFIDDDELEFIYDRLELTKYK 45
OY      448 GYOVPAPLESVLLQHPNIPAGVAGVDDPIAGELPGVAVVLYKKGSKMTEKEVWDVYASO 50
Db      458 GFOVAPAELESLLIGHPEINDVAVVAKKEBDAAGEVPAFVYKRSQDSNISSEDELKOPFSKO 51
OY      508 VSNAKRLBGAGRVFDEVYKGLTKLIDGKALR 538
Db      518 VVYFKRI--NKYFETDSIPKADSGILKKDR 547

```

Search completed: November 29, 2001, 02:00:07
Job time: 377 sec

Query Match	25.2%;	Score 710;	DB 1;	Length 556;
Best Local Similarity	34.6%;	Pred. No. 5.5e-42;		
Matches 177;	Conservative 92;	Mismatches 204;	Indels 38;	Gaps

